GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 14:06:06 ; Search time 17.59 Seconds
(without alignments)
1001.557 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2282	99.7	455	1	AROA_AGRSP	Q9r4e4 agrobacteri
2	1900.5	83.1	449	1	AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1	AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1	AROA_SYNY3	Q59975 synechocyst
5	806.5	35.2	430	1	AROA_LACLA	Q9ceu0 lactococcus
6	806	35.2	443	1	AROA_BACNO	Q46550 bacteroides
7	803.5	35.1	427	1	AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1	AROA_LACLC	P43905 lactococcus
9	785	34.3	428	1	AROA_BACSU	P20691 bacillus su
10	766.5	33.5	431	1	AROA_AQUAE	O67494 aquifex aeo
11	695	30.4	410	1	AROA_THEMA	Q9wyi0 thermotoga
12	590	25.8	428	1	AROA_CAMJE	P52312 campylobact
13	577	25.2	430	1	AROA_STAAU	Q05615 staphylococ
14	521.5	22.8	429	1	AROA_HELPJ	Q9zkf7 helicobacte
15	518.5	22.7	429	1	AROA_HELPY	P56197 helicobacte
16	496	21.7	419	1	AROA_METTH	O26860 methanobact
17	459.5	20.1	429	1	AROA_METJA	Q57925 methanococc
18	408	17.8	439	1	AROA_DEIRA	Q9rvd3 deinococcus
19	403	17.6	416	1	AROA_ARCFU	O28775 archaeoglob
20	402.5	17.6	427	1	AROA_AERPE	Q9yek9 aeropyrum p
21	402.5	17.6	427	1	AROA_YEREN	P19688 yersinia en
22	395.5	17.3	428	1	AROA_YERPE	Q60112 yersinia pe
23	387.5	16.9	432	1	AROA_HAESO	P52310 haemophilus
24	378	16.5	427	1	AROA_ECOLI	P07638 escherichia
25	376	16.4	427	1	AROA_SHISO	Q9zff7 shigella so
26	373.5	16.3	410	1	AROA_PYRAB	Q9v1h1 pyrococcus
27	373	16.3	432	1	AROA_PASHA	P54220 pasteurella
28	372	16.3	427	1	AROA_KLEPN	P24497 klebsiella
29	372	16.3	440	1	AROA_PASMU	Q04570 pasteurella
30	371	16.2	427	1	AROA_SHIDY	O87006 shigella dy
31	365.5	16.0	463	1	AROA_BURPS	P39915 burkholderi
32	365	16.0	428	1	AROA_EDWIC	Q9x4h2 edwardsiell
33	363.5	15.9	432	1	AROA_HAEIN	Q03421 haemophilus
34	357	15.6	427	1	AROA_SALTY	P07637 salmonella
35	348	15.2	427	1	AROA_SALGL	P22299 salmonella
36	347	15.2	427	1	AROA_SALTI	P19786 salmonella
37	323.5	14.1	516	1	AROA_PETHY	P11043 petunia hyb
38	321.5	14.1	1588	1	ARO1_YEAST	P08566 s pentafunc
39	317	13.9	441	1	AROA_CHLMU	Q9pk28 chlamydia m
40	315.5	13.8	518	1	ARO1_TOBAC	P23981 nicotiana t
41	314.5	13.7	440	1	AROA_CHLTR	O84371 chlamydia t
42	313	13.7	428	1	AROA_BUCAP	Q59178 buchnera ap
43	310.5	13.6	520	1	AROA_LYCES	P10748 lycopersico
44	308.5	13.5	1573	1	ARO1_SCHPO	Q9p7r0 s pentafunc
45	306.5	13.4	445	1	AROA_CHLPN	Q9z6m0 chlamydia p

ALIGNMENTS

RESULT 1
AROA_AGRSP
ID AROA_AGRSP STANDARD; PRT; 455 AA.
AC Q9R4E4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Agrobacterium sp. (strain CP4).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=361;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.
RA Barry G.F., Kishore G.M., Padgette S.R., Stallings W.C.;
RT "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL Patent number US5633435, 27-MAY-1997.
RN [2]
RP SEQUENCE OF 1-15.
RX MEDLINE=96182485; PubMed=8598558;
RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,
RA Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,
RA Fuchs R.L., Padgette S.R.;
RT "The expressed protein in glyphosate-tolerant soybean, 5-
RT enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp.
RT strain CP4, is rapidly digested in vitro and is not toxic to acutely
RT gavaged mice.";
RL J. Nutr. 126:728-740(1996).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in
CC glyphosate-tolerant soybean, canola, cotton and maize by Monsanto.
CC Developed to provide new weed-control options for farmers.
CC Expression of this protein in plants imparts high levels of
CC glyphosate tolerance.
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR InterPro; IPR001986; EPSP_syntase.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance;
KW Genetically modified food.
FT CONFLICT 2 2 S -> L (IN REF. 2).
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;

Query Match 99.7%; Score 2282; DB 1; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.1e-134;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   2
AROA_PSES2
ID   AROA_PSES2     STANDARD;      PRT;   449 AA.
AC   P56952;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Pseudomonas sp. (strain PG2982), and
OS   Achromobacter sp. (strain LBAA).
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=308, 129026;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 2-16.
RA   Barry G.F., Kishore G.M., Padgette S.R., Stallings W.C.;
RT   "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL   Patent number US5633435, 27-MAY-1997.
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- MISCELLANEOUS: RESISTANT TO THE ANTIBIOTIC GLYPHOSATE.
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Herbicide resistance.
SQ   SEQUENCE   449 AA;  47297 MW;  447F213EECCAEFC1 CRC64;
```

```
 Query Match           83.1%;  Score 1900.5;  DB 1;  Length 449;
 Best Local Similarity  82.9%;  Pred. No. 4.8e-111;
 Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;
```

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: |||||:|:|| |||| ||||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   | |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447
```

```
RESULT   3
AROA_BACHD
ID   AROA_BACHD     STANDARD;      PRT;   431 AA.
AC   Q9KCA6;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR AROE OR BH1667.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
CC       BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; AP001512; BAB05386.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   431 AA;  45485 MW;  12F4FFBE7BA0743D CRC64;


  Query Match             38.6%;  Score 884;  DB 1;  Length 431;
  Best Local Similarity   44.2%;  Pred. No. 5e-48;
  Matches  188;  Conservative   72;  Mismatches  153;  Indels   12;  Gaps    5;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           : || ||:::|||||||||: ||| || | | : | | | | ::|    | :|  | :
Db      10 AKGLKGTIKVPGDKSISHRAVMFGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |  ||  |   || ||: |  || :|::    | |  ||| |: |||| ||
Db      70 ERVTVKGKGWDGLREPSDILDVGNSGTTTRLILGILSTLPFHSVIIGDESIGKRPMKRVT 129

Qy     135 NPLREMGVQVKSED-GDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || |:   | |:  |:::|| :    | :  |:||||:|||:||||   | |:|
Db     130 EPLKSMGAQIDGRDHGNLTPLSIRGGQL-KGIDFHSPVASAQMKSAILLAGLRAEGKTSV 188

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           ||  ||||||:||: || |:    : ||: |: :||   |||| : |||| || || ||
Db     189 TEPAKTRDHTERMLEAFGVNI----EKDGL-TVSIEGGQMLTGQHVVVPGDISSAAFFLV 243

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
           |  :|| | :|: || :|||| |::  |::||| : : | |: ||| |||| : :| |:|
Db     244 AGAMVPHSRITLTNVGINPTRAGILEVLKQMGATLAMENERVQGGEPVADLTIETSVLQG 303

Qy     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373
           | :  |  | :||| ||:|| |  | | ||:   |||:|||::|:  | : |   |
Db     304 VEIGGDIIPRLIDEIPIIAVLATQASGRTVIKDAEELKVKETNRIDTVVSELTKLGASIH 363

Qy     374 EGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSF 433
           :  ::: |    ||     |  |::| |||| |:  :  |::| ||||:    || |:
Db     364 ATDDGMIIEGPTPLKG-----GVTVSSHGDHRIGMAMAIAALLAEKPVTVEGTEAIAVSY 418

Qy     434 PEFMD 438
           | | |
Db     419 PSFFD 423


RESULT   4
_SYNY3
     AROA_SYNY3     STANDARD;      PRT;   447 AA.
AC   Q59975; Q59974;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR SLR0444.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94299161; PubMed=8026753;
RA   dalla Chiesa M., Mayes S.R., Maskell D.S., Nixon P.J., Barber J.;
RT   "An aroA homologue from Synechocystis sp. PCC 6803.";
RL   Gene 144:145-146(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
RN   [3]
RP   SEQUENCE OF 103-194 FROM N.A.
RX   MEDLINE=93307506; PubMed=7686511;
RA   Mayes S.R., dalla Chiesa M., Zhang Z., Barber J.;
RT   "The genes aroA and trnQ are located upstream of psbO in the
RT   chromosome of Synechocystis 6803.";
RL   FEBS Lett. 325:255-261(1993).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75325; CAA53074.1; -.
DR   EMBL; D90914; BAA18477.1; -.
DR   EMBL; X72784; CAA51291.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   447 AA;  47046 MW;  2F3C8AD26B5A7BCE CRC64;


  Query Match             37.9%;  Score 867.5;  DB 1;  Length 447;
  Best Local Similarity   48.1%;  Pred. No. 5.5e-47;
  Matches  203;  Conservative   62;  Mismatches  144;  Indels   13;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRK-EGDT 76
           |:| :|:|||||||||: | | :|:||| | ||| |||  :|    :|||| | :   :
Db      26 LTGRLRVPGDKSISHRALMLGAIATGETIIEGLLLGEDPRSTAHCFRAMGAEISELNSEK 85

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLV-GVYDFDSTFIGDASLTKRPMGRVLN 135
            |: | | | |  |   || ||: |  || :||: |  |   |  || ||  ||| ||:
Db      86 IIVQGRGLGQLQEPSTVLDAGNSGTTMRLMLGLLAGQKDCLFTVTGDDSLRHRPMSRVIQ 145

Qy     136 PLREMGVQV-KSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
           ||::|| ::    :|   |: ::| :   || |  |:|||||| :|||| | | |||
Db     146 PLQQMGAKIWARSNGKFAPLAVQGSQL-KPIHYHSPIASAQVKSCLLLAGLTTEGDTTVT 204

Qy     195 EPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVA 254
           || ::|||:|:||| ||| ||::       :: : |   |||| : |||| || || |||
Db     205 EPALSRDHSERMLQAFGAKLTID---PVTHSVTVHGPAHLTGQRVVVPGDISSAAFWLVA 261

Qy     255 ALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGV 314
           | ::|||:: : || :||||||::  | :|||||   | ||  || |||||||:| |:|
Db     262 ASILPGSELLVENVGINPTRTGVLEVLAQMGADITPENERLVTGEPVADLRVRASHLQGC 321

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
           |   :  | :||| |||||||||||| | :    ||||||||||:|:|: |   |   |
Db     322 TFGGEIIPRLIDEIPILAVAAAFAEGTTRIEDAAELRVKESDRLAAIASELGKMGAKVTE 381

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFP 434
           :  | ::|   | |    || | :  ||||||:  :  | |     :: |   | |:|
Db     382 FDDGLEIQG---GSPL---QGAEVDSLTDHRIAMALAIAALGSGGQTIINRAEAAAISYP 435

Qy     435 EF 436
```

```
Db    436 EF 437 .


RESULT   5
AROA_LACLA
ID   AROA_LACLA     STANDARD;      PRT;   430 AA.
AC   Q9CEU0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR LL1744.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006404; AAK05842.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   430 AA;  45958 MW;  CB216F07AA4EE799 CRC64;


  Query Match             35.2%;  Score 806.5;  DB 1;  Length 430;
  Best Local Similarity   41.9%;  Pred. No. 3.2e-43;
  Matches  184;  Conservative   77;  Mismatches  157;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| ||||::| :| :||| | :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTVIYDILRGEDVLSTIEAFRAMGVEIEDKG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
             :  : | |   | |||  || ||: |  ||  |::    |::|  || ||:|||| ||
Db      67 EVITVHGKGISELKAPEKALDMGNSGTSTRLLSGILAGLPFETTLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|      | | :|:|||||||||: | |   |:| |
Db     127 TPLQLMGAEITGQTDKVKLPMTIKGSTHLKAIDYVLPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           :|   || |||:||  ||  | |       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGELKVSD-----KTILVPGGQKLVGQKVVVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           |||:|  |:: : || :| ||||:|  :| ||  :|::       :|     | |:|::|
Db     242 AALVVENSELILENVGVNETRTGIIEVIQAMGGQLEILE------QDNVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
           |||  :  |  | :||| ||:|: |  |:| |::    ||:|||:||:: ||| |   |
Db     296 QLKGTEISGDLIPRLIDELPIIALLATQAQGQTIIRDAAELKVKETDRIAVVANALNSMG 355

Qy     370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428
                :  ::::|   |  | :|   :: |  |||| |   :  |: :|  : :: |
Db     356 AKIQPTDDGMIIQG---GTKL-HAPENSINTLGDHRIGMMAAIAALLVKNGEIELERAEA 411

Qy     429 IATSFPEFMDLMAGLGAKI 447
           | ||:| | | :  |   |
Db     412 IQTSYPSFFDDLEQLSENI 430


RESULT   6
AROA_BACNO
ID   AROA_BACNO     STANDARD;      PRT;   443 AA.
AC   Q46550;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Bacteroides nodosus (Dichelobacter nodosus).
OC   Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OC   Dichelobacter.
OX   NCBI_TaxID=870;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VCS1001;
RX   MEDLINE=94320795; PubMed=8045432;
RA   Alm R.A., Dalrymple B.P., Mattick J.S.;
RT   "Sequencing and expression of the aroA gene from Dichelobacter
RT   nodosus.";
RL   Gene 145:97-101(1994).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z29339; CAA82544.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   443 AA;  47444 MW;  BE2243277ADEFBD5 CRC64;


  Query Match             35.2%;  Score 806;  DB 1;  Length 443;
  Best Local Similarity   42.2%;  Pred. No. 3.6e-43;
  Matches  185;  Conservative   74;  Mismatches  169;  Indels   10;  Gaps    3;

Qy      16 SGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75
           | ||| : | ||||:|||: :   || |:| | | |   | : | :|::|:|  |::| :
Db      12 SALSGEITICGDKSMSHRALLLAALAEGQTEIRGFLACADCLATRQALRALGVDIQREKE 71
```

```
Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db      72 IVTIRGVGFLGLQPPKAPLNMQNSGTSMRLLAGILAAQRFESVLCGDESLEKRPMQRIIT 131

Qy     136 PLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
Db     132 PLVQMGAKIVSHSNFTAPLHISG-RPLTGIDYALPLPSAQLKSCLILAGLLADGTTRLHT 190

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
Db     191 CGISRDHTERMLPLFGGALEIKKE-----QIIVTGGQKLHGCVLDIVGDLSAAAFFMVAA 245

Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
Db     246 LIAPRAEVVIRNVGINPTRAAIITLLQKMGGRIELHHQRFWGAEPVADIVVYHSKLRGIT 305

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
Db     306 VAPEWIANAIDELPIFFIAAACAEGTTFVGNLSELRVKESDRLAAMAQNLQTLGVACDVG 365

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPE 435
Db     366 ADFIHIYGRSDRQFL----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMPQ 421

Qy     436 FMDLMAGLGAKIELSDTK 453
Db     422 FRDFAAAIGMNVGEKDAK 439
```

```
RESULT   7
AROA_STRPN
ID   AROA_STRPN     STANDARD;      PRT;   427 AA.
AC   Q9S400;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR SP1371.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=NCIMB 40794 / 0100993;
RX   MEDLINE=20069365; PubMed=10601870;
RA   Du W., Wallis N.G., Mazzulla M.J., Chalker A.F., Zhang L., Liu W.-S.,
RA   Kallender H., Payne D.J.;
RT   "Characterization of Streptococcus pneumoniae 5-enolpyruvylshikimate
RT   3-phosphate synthase and its activation by univalent cations.";
RL   Eur. J. Biochem. 267:222-227(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF169483; AAD45819.1; -.
DR   EMBL; AE007434; AAK75469.1; -.
DR   TIGR; SP1371; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT   CONFLICT     13     13       S -> I (IN REF. 1).
FT   CONFLICT     71     71       I -> V (IN REF. 1).
FT   CONFLICT    201    201       K -> Q (IN REF. 1).
FT   CONFLICT    303    303       G -> C (IN REF. 1).
SQ   SEQUENCE   427 AA;  45766 MW;  45CE6F4D0D1C7B70 CRC64;
```

```
  Query Match             35.1%;  Score 803.5;  DB 1;  Length 427;
  Best Local Similarity   42.8%;  Pred. No. 4.9e-43;
  Matches  184;  Conservative   72;  Mismatches  155;  Indels   19;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
Db      10 LHGSIRVPGDKSISHRSIIFGSLAEGETKVYDILRGEDVLSTMQVFRDLGVEIEDKDGVI 69

Qy      78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
Db      70 TIQGVGMAGLKAPQNALNMGNSGTSIRLISGVLAGADFEVEMFGDDSLSKRPMDRVTLPL 129

Qy     138 REMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196
Db     130 KKMGVSISGQTERDLPPLRLKGTKNLRPIHYELPIASAQVKSALMFAALQAKGESVIIEK 189

Qy     197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
Db     190 EYTRNHTEDMLKQFGGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISSAAFWLVAGL 244

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKGV 314
Db     245 IAPNSRLVLQNVGINETRTGIIDVIRAMGGKLEITEIDPVAKS----ATLIVESSDLKGT 300

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
Db     301 EIGGALIPRLIDELPIIALLATQAQGVTVIKDAEELKVKETDRIQVVADALNSMGADITP 360

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSF 433
Db     361 TADGMIIKGK------SALHGARVNTFGDHRIGMMTAIAALLVADGEVELDRAEAINTSY 414

Qy     434 PEFMDLMAGL 443
Db     415 PSFFDDLESL 424
```

```
RESULT   8
AROA_LACLC
ID   AROA_LACLC     STANDARD;      PRT;   430 AA.
AC   P43905;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
```

```
OS   Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1359;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MG1363 / F15876;
RX   MEDLINE=95124293; PubMed=7823907;
RA   Griffin H.G., Gasson M.J.;
RT   "Genetic aspects of aromatic amino acid biosynthesis in Lactococcus
RT   lactis.";
RL   Mol. Gen. Genet. 246:119-127(1995).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78413; CAA55180.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   430 AA;  45804 MW;  C5A197A49072C9D5 CRC64;
```

```
  Query Match           35.1%;  Score 802.5;  DB 1;  Length 430;
  Best Local Similarity  41.2%;  Pred. No. 5.7e-43;
  Matches  181;  Conservative   79;  Mismatches  158;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| ||||::| :| :|:|  |  :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTIIHDILRGEDVLSTIEAFRALGVEIEDDG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
              : | |   |  ||  || ||: |  ||  |::    |::|  || ||:|||| ||
Db      67 QVITVHGQGISKLKEPEKALDMGNSGTSTRLLSGILAGLPFEATLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|      | | :|:|||||||||: | |   |:| |
Db     127 TPLQMMGAEIVGQTDKVKLPMTIKGSAHLKAIDYILPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           :|   || |||:||  ||  :||       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGEITVSD-----KTILVPGGQKLLGQEVTVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           | |:|  | : : || :| |||||:  :| ||  :|::       :|      | |:|::|
Db     242 AGLVVENSGLILENVGINETRTGILEVIQAMGGQLEILE------QDEVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
           |||  :  |  | :||| ||:|: |  ||| |::    ||:|||:||:: ||: |   |
Db     296 QLKGTEISGDLIPRLIDELPIIALLATQAEGKTIIRDAAELKVKETDRIAVVADALNSMG 355

Qy     370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428
           : :  :  ::::|   |  | :|   |: |  |||| |   :  |: ||  : :: |
Db     356 ANIEPTDDGMIIQG---GTKL-HAPENAINTLGDHRIGMMVAIAALLVENGEIELERAEA 411

Qy     429 IATSFPEFMDLMAGLGAKI 447
             | ||:| | | :  |   :
Db     412 IQTSYPSFFDDLEKLSGNL 430
```

```
RESULT   9
AROA_BACSU
ID   AROA_BACSU     STANDARD;      PRT;   428 AA.
AC   P20691;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROE.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191994; PubMed=3106153;
RA   Henner D.J., Band L., Flaggs G., Chen E.;
RT   "The organization and nucleotide sequence of the Bacillus subtilis
RT   hisH, tyrA and aroE genes.";
RL   Gene 49:147-152(1986).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80245; AAA20869.1; -.
DR   EMBL; Z99115; CAB14176.1; -.
DR   PIR; C26532; C26532.
DR   SubtiList; BG10294; aroE.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   428 AA;  45240 MW;  DE3F7B96E761CB40 CRC64;
```

```
  Query Match           34.3%;  Score 785;  DB 1;  Length 428;
  Best Local Similarity  41.0%;  Pred. No. 6.9e-42;
  Matches  178;  Conservative   68;  Mismatches  176;  Indels   12;  Gaps    4;

Qy      14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73
           |   | | : |||||||||| ||| ||:| | :   | | | ::|    : ||  | :
Db       5 KVQTLHGEIHIPGDKSISHRSVMFGALAAGTTTVKNFLPGADCLSTIDCFRKMGVHIEQS 64

Qy      74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133
              :| | |   |  ||: || ||: |  || :|::    | |   || |: |||| ||
Db      65 SSDVVIHGKGIDALKEPESLLDVGNSGTTIRLMLGILAGRPFYSAVAGDESIAKRPMKRV 124

Qy     134 LNPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192
             ||::|| ::     |:  |::: |      | |  |:||||:|||||||||   | ||
Db     125 TEPLKKMGAKIDGRAGGEFTPLSVSGASL-KGIDYVSPVASAQIKSAVLLAGLQAEGTTT 183

Qy     193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252
           | ||  :|||||:||  ||  |     ::   :: : |  |||   | |||| || || |
Db     184 VTEPHKSRDHTERMLSAFGVKL-----SEDQTSVSIAGGQKLTAADIFVPGDISSAAFFL 238

Qy     253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLK 312
```

```
           |  :|| | : : || :||||||:|  || ||| :|:     :| |   || : :|:||
Db    239 AAGAMVPNSRIVLKNVGLNPTRTGIIDVLQNMGAKLEIKPSADSGAEPYGDLIIETSSLK 298

Qy    313 GVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDC 372
          | :  |  | :||| ||:|: |  ||| ||:    ||:|||::|:  | : |:  | :
Db    299 AVEIGGDIIPRLIDEIPIIALLATQAEGTTVIKDAAELKVKETNRIDTVVSELRKLGAEI 358

Qy    373 DEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATS 432
          :     : | |:   ||     ||||::| |||| |   :   ::| |: ::    |  |
Db    359 EPTADGMKVYGKQTLKG-----GAAVSSHGDHRIGMMLGIASCITEEPIEIEHTDAIHVS 413

Qy    433 FPEFMDLMAGLGAK 446
          :| | : :  |  |
Db    414 YPTFFEHLNKLSKK 427


RESULT  10
AROA_AQUAE
ID   AROA_AQUAE     STANDARD;      PRT;   431 AA.
AC   O67494;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR AQ_1536.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000744; AAC07443.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   431 AA;  47793 MW;  EF842512EBE41D2A CRC64;


  Query Match             33.5%;  Score 766.5;  DB 1;  Length 431;
  Best Local Similarity   40.7%;  Pred. No. 9.7e-41;
  Matches  174;  Conservative   73;  Mismatches  169;  Indels   11;  Gaps    6;

Qy     14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73
          |   : | :|:| ||||:||:|: | |||||| :   |   | : | : ::|:  ::|:
Db      6 KIKRVKGELRVPSDKSITHRAFILGALASGETLVRKPLISGDTLATLEILKAIRTKVREG 65

Qy     74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133
             :  :|:|  |   | |   ||  |: |  |:  |::    | |   || ||  ||| ||
Db     66 KEEVLIEG-RNYTFLEPHDVLDAKNSGTTARIMSGVLSTQPFFSVLTGDESLKNRPMLRV 124

Qy    134 LNPLREMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192
          : |||||| ::   |:|:::|: :||      |:|    :|||||||:|||||   |:|
Db    125 VEPLREMGAKIDGREEGNKLPIAIRGGNL-KGISYFNKKSSAQVKSALLLAGLRAEGMTE 183

Qy    193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252
          |:|| ::|||||:||: ||| :    :  |   ::::|  :| |  :  | |||| |:
Db    184 VVEPYLSRDHTERMLKLFGAEVITIPEERG-HIVKIKGGQELQGTEVYCPADPSSAAYFA 242

Qy    253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRS-STL 311
            | |  :: :  ||:|||| |    | ||| ||   | |    | :||| ||    |
Db    243 ALATLAPEGEIRLKEVLLNPTRDGFYRKLIEMGGDISFENYRELSNEPMADLVVRPVDNL 302

Qy    312 KGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVD 371
          | | |  :  |::||| |||||  |||:| : : | :||| |||||: |:   |:  ||
Db    303 KPVKVSPEEVPTLIDEIPILAVLMAFADGVSEVKGAKELRYKESDRIKAIVTNLRKLGVQ 362

Qy    372 CDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIAT 431
            :| |    : |  : |      |  : |  ||||||:| |:||| |   | :|    :
Db    363 VEEFEDGFAIHGTKEIK------GGVIETFKDHRIAMAFAVLGLVVEEEVIIDHPECVTV 416

Qy    432 SFPEFMD 438
          |:||| :
Db    417 SYPEFWE 423


RESULT  11
AROA_THEMA
ID   AROA_THEMA     STANDARD;      PRT;   410 AA.
AC   Q9WYI0;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR TM0345.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001715; AAD35431.1; -.
DR   TIGR; TM0345; -.
```

```
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   410 AA;  45341 MW;  DC9F7D44792CA69F CRC64;


  Query Match           30.4%;  Score 695;  DB 1;  Length 410;
  Best Local Similarity  40.5%;  Pred. No. 2.5e-36;
  Matches  170;  Conservative   67;  Mismatches  161;  Indels   22;  Gaps   10;

Qy      24 IPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVG 83
           :| ||||:||: :   ||  |: :  ||   |   |   :: :| |   ||| |    |
Db       4 VPPDKSITHRALILSALAETESTLYNLLRCLDTERTHDILEKLGTRF--EGD-WEKMKVF 60

Qy      84 NGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQ 143
             |  ||  ||: |  ||  |::  |:  :   || ||::||| ||: ||  || :
Db      61 PKPFAEPIEPLFCGNSGTTRLMSGVLASYEMFTVLYGDPSLSRRPMRRVIEPLEMMGAR 120

Qy     144 VKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIMTRDHT 203
             :   : ||: ::|   : |:|: |:|||||||||||||   | | ||||  :||||
Db     121 FMARQNNYLPMAIKGNHL-SGISYKTPVASAQVKSAVLLAGLRASGRTIVIEPAKSRDHT 179

Qy     204 EKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDV 263
           |:||:    || |  : :| | : ||     |  : |||| || || :|   : | : :
Db     180 ERMLK----NLGVPVEVEGTRVV-LE-PATFRGFTMKVPGDISSAAFFVVLGAIHPNARI 233

Qy     264 TILNVLMNPTRTGLILTLQEMGADI--EVINPRLAGGEDVADLRVRSS-TLKGVTVPEDR 320
           |: :| :|||||||:  :: |||::  |:    |   | :  :|| :|  |||| |||
Db     234 TVTDVGLNPTRTGLLEVMKLMGANLEWEITEENL---EPIGTVRVETSPNLKGVVVPEHL 290

Qy     321 APSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEGETSLV 380
             | |||| |::|:   |||| ||:   |||| |||||:  :    |  ||: :| :
Db     291 VPLMIDELPLVALLGVFAEGETVVRNAEELRKKESDRIRVLVENFKRLGVEIEEFKDGFK 350

Qy     381 VRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPEFMDLM 440
           : {:   |    | :|   |||:|| | : |||||  | | |   :| ||| | :|:
Db     351 IVGKQSIK------GGSVDPEGDHRMAMLFSIAGLVSEEGVDVKDHECVAVSFPNFYELL 404


RESULT  12
AROA_CAMJE
ID   AROA_CAMJE     STANDARD;      PRT;   428 AA.
AC   P52312; Q9PP36;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR CJ0895C.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=81116;
RX   MEDLINE=97128776; PubMed=8973316;
RA   Woesten M.M.S.M., Dubbink V.H.J., van der Zeijst B.A.M.;
RT   "The aroA gene of Campylobacter jejuni.";
RL   Gene 181:109-112(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X89371; CAA61554.1; -.
DR   EMBL; AL139076; CAB73153.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT   CONFLICT    152    152       K -> N (IN REF. 1).
FT   CONFLICT    158    158       F -> Y (IN REF. 1).
FT   CONFLICT    179    179       D -> N (IN REF. 1).
FT   CONFLICT    183    183       T -> A (IN REF. 1).
FT   CONFLICT    217    217       N -> S (IN REF. 1).
FT   CONFLICT    241    241       V -> A (IN REF. 1).
SQ   SEQUENCE   428 AA;  47288 MW;  465E2B50F39FC5AB CRC64;


  Query Match           25.8%;  Score 590;  DB 1;  Length 428;
  Best Local Similarity  34.0%;  Pred. No. 8.5e-30;
  Matches  146;  Conservative   94;  Mismatches  167;  Indels   22;  Gaps   11;

Qy      24 IPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVG 83
           |  |||||||  :|  |   | :    |  :| :|| : :: :||:| ::     :  :
Db      17 IAADKSISHRFAIFSLLTQEENKAQNYLLAQDTLNTLEIIKNLGAKIEQKDS--CVKIIP 74

Qy      84 NGGLLAPEAPLDFGNAATGCRLTMG-LVGVYDFDSTFI--GDASLTKRPMGRVLNPLREM 140
             :|:|   || ||: |  || :| | |:  |   |:  ||  |  ||| |:  || ::
Db      75 PKEILSPNCILDCGNSGTAMRLMIGFLAGISGF---FVLSGDKYLNNRPMRRISKPLTQI 131

Qy     141 GVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIMT 199
           | ::   :   | |: : | |      ::  ::|||||:|::|:      : |  |  ::
Db     132 GARIYGRNEANLAPLCIEGQKLKA-FNFKSEISSAQVKTAMILSAFRADNVCTFSEISLS 190

Qy     200 RDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVP 259
           |:|:| ||:   | : |  |   :   |  :  | | | :| ||||  : ::||:::|
Db     191 RNHSENMLKAMKAPIRVSNDGLSLEINPL--KKPLKAQNIIIPNDPSSAFYFVLAAIILP 248

Qy     260 GSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPED 319
             | : |:|:|||       ||:||| :|:   :    | : ::|| || | |: | :|
Db     249 KSQIILKNILLNPTRIEAYKILQKMGAKLEMTITQ-NDFETIGEIRVESSKLNGIEV-KD 306

Qy     320 RAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEGETSL 379
             :||| | ||:| | |:| : :   :|||||||||:: :   ||| ||:  | :
Db     307 NIAWLIDEAPALAIAFALAKGKSSLINAKELRVKESDRIAVMVENLKLCGVEARELDDGF 366

Qy     380 VVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPEFMDL 439
           : |  : |       : : :: |||||||| ::||:    : :||:  | |||| |:::
Db     367 EIEGGCELK------SSKIKSYGDHRIAMSFAILGLLC--GIEIDDSDCIKTSFPNFIEI 418

Qy     440 MAGLGAKIE 448
           :: |||:|:
Db     419 LSNLGARID 427


RESULT  13
```

```
AROA_STAAU
ID   AROA_STAAU     STANDARD;      PRT;   430 AA.
AC   Q05615;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 8325-4;
RX   MEDLINE=93381456; PubMed=8371108;
RA   O'Connell C.M., Pattee P., Foster T.J.;
RT   "Sequence and mapping of the aroA gene of Staphylococcus aureus
RT   8325-4.";
RL   J. Gen. Microbiol. 139:1449-1460(1993).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L05004; AAA71897.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   430 AA;  46852 MW;  E5FC878EA1C23C20 CRC64;


  Query Match           25.2%;  Score 577;  DB 1;  Length 430;
  Best Local Similarity   33.6%;  Pred. No. 5.5e-29;
  Matches  144;  Conservative   82;  Mismatches  179;  Indels   24;  Gaps    8;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           |  | | : :|||||::||: |   || | : |   | |||   |    : :|  |::: 
Db      10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGVSTIYKPLLGEDCRRTMDIFRHLGVEIKEDD 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :  ::   |   : |   | ||: |  ||  ||:     :|   || |: |||| |||
Db      70 EKLVVTSPGY-QVNTPHQVLYTGNSGTTTRLLAGLLSGLGNESVLSGDVSIGKRPMDRVL 128

Qy     135 NPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
            ||: |   ::  : :  |: :: |     | |:: :|||||||:| | | :   | :
Db     129 RPLKLMDANIEGIEDNYTPLIIK-PSVIKGINYQMEVASAQVKSAILFASLFSKEPTIIK 187

Qy     195 EPIMTRDHTEKMLQGF-----GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249
           |  ::|:||| | : |       |:: |  : :| |:             |||| || |
Db     188 ELDVSRNHTETMFKHFNIPIEAEGLSINTTPEAIRYIKPAD--------FHVPGDISSAA 239

Qy     250 FPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVR-S 308
           | :|||: |||||| || :| ||:|:|  :::|| :|:: | :  | |  | :|:: :
Db     240 FFIVAALITPGSDVTIHNVGINQTRSGIIDIVEKMGGNIQLFN-QTTGAEPTASIRIQYT 298

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
             |: :|:  :  |  ||| |::|:    | | : :   |||:|||::|:   |: | |

Db     299 PMLQPITIEGELVPKAIDELPVIALLCTQAVGTSTIKDAEELKVKETNRIDTTADMLNLL 358

Qy     369 GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATM 428
           | :       |::      :   ||:        |||| |   |  ::|  || :
Db     359 GFELQPTNDGLIIH---PSEFKTNATDILT----DHRIGMMLAVACVLSSEPVKIKQFDA 411

Qy     429 IATSFPEFM 437
           :  ||| |:
Db     412 VNVSFPGFL 420


RESULT  14
AROA_HELPJ
ID   AROA_HELPJ     STANDARD;      PRT;   429 AA.
AC   Q9ZKF7;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR JHP0980.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001527; AAD06557.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   429 AA;  47167 MW;  92724C4A25752741 CRC64;


  Query Match           22.8%;  Score 521.5;  DB 1;  Length 429;
  Best Local Similarity   33.6%;  Pred. No. 1.5e-25;
  Matches  145;  Conservative   73;  Mismatches  166;  Indels   47;  Gaps   13;

Qy      27 DKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG--------DTWI 78
           |||:|||: :|  ||     :   | ||| ::: :  | :||::           | |
Db      10 DKSLSHRAVIFSLLAQKPCFVRNFLMGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTTI 69

Qy      79 IDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLR 138
           :          |   |:  |: | ||  ||:          || ||  ||| |:: ||:
Db      70 KE---------PNKILNCNNSGTSMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK 120
```

```
Qy    139 EMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPI 197
          | ::   ||    |: : |         |  |:||||||| :|: |   ||:   |
Db    121 AFGAKILGREDNHFAPLAIVGGPLKA-CDYESPIASAQVKSAFILSALQAQGISAYKESE 179

Qy    198 MTRDHTEKMLQGFGANLTVETDADGVRTIR-LEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
          ::|:||| ||:  |||:    : ||| |  ||    |      :  |||| |  :|
Db    180 LSRNHTEIMLKSLGANI---QNQDGVLKISPLE--KPLESFDFTIANDPSSAFFLALACA 234

Qy    257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIE-VINPRLAGGEDVADLRVRSSTLKGVT 315
          : | | : : |||:|||      |::||| || ||  :    | : |: :  : || ::
Db    235 ITPKSRLLLKNVLLNPTRIEAFEVLKKMGAHIEYVIQSK--DLEVIGDIYIEHAPLKAIS 292

Qy    316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
          : :: | |:||| | |::|  ||:| :::   ::|| |||||: || :  |  |::|:|
Db    293 IDQNIA-SLIDEIPALSIAMLFAKGKSMVRNAKDLRAKESDRIKAVVSNFKALGIECEEF 351

Qy    376 ETSLVVRGRPDGKGLGNAS----------GAAVATHLDHRIAMSFLVMGLVSENPVTVDD 425
          |    :      :|||:||            : :  |||||||| |: |    |: :|:
Db    352 EDGFYI------EGLGDASQLKQHFSKIKPPIIKSFNDHRIAMSFAVLTLAL--PLEIDN 403

Qy    426 ATMIATSFPEF 436
            ||| |
Db    404 LECANISFPTF 414
```

```
RESULT  15
AROA_HELPY
ID   AROA_HELPY     STANDARD;      PRT;   429 AA.
AC   P56197;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR HP0401.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000556; AAD07470.1; -.
DR   TIGR; HP0401; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   429 AA;  47240 MW;  19545753E081FDAE CRC64;
```

Query Match 22.7%; Score 518.5; DB 1; Length 429;
Best Local Similarity 33.5%; Pred. No. 2.3e-25;
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

```
Qy     27 DKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG--------DTWI 78
          |||:|||: :|  ||      :   | ||| ::: :  | :||::             | |
Db     10 DKSLSHRAVIFSLLAQKPCFVRNFLMGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTTI 69

Qy     79 IDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLR 138
          :       |   |:  |: |  ||  ||:          || ||  ||| |:: ||:
Db     70 KE---------PNKILNCNNSGTTMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK 120

Qy    139 EMGVQVKS-EDGDRLPVTLRGPKTPTPIT---YRVPMASAQVKSAVLLAGLNTPGITTVI 194
          | ::   ||    |: : |    :|:    |  |:||||||| :|: |   | :|
Db    121 AFGAKILGREDNHFAPLVILG----SPLKACHYESPIASAQVKSAFILSALQAQGASTYK 176

Qy    195 EPIMTRDHTEKMLQGFGANLTVETDADGVRTIR-LEGRGKLTGQVIDVPGDPSSTAFPLV 253
          |  ::|:||| ||:  ||::    : |||  |  ||     |      :  ||||  | :
Db    177 ESELSRNHTEIMLKSLGADI---HNQDGVLKISPLE--KPLEAFDFTIANDPSSAFFFAL 231

Qy    254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIE--VINPRLAGGEDVADLRVRSSTL 311
          |  : | | : : |||:||||      |::||| ||  : :  |   | : |: |  : |
Db    232 ACAITPKSRLLLKNVLLNPTRIEAFEVLKKMGASIEYAIQSKDL---EMIGDIYVEHAPL 288

Qy    312 KGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVD 371
          | : : :: | |:||| | |::|  ||:| :::   ::|| |||||: || :  |  |::
Db    289 KAINIDQNIA-SLIDEIPALSIAMLFAKGKSMVKNAKDLRAKESDRIKAVVSNFKALGIE 347

Qy    372 CDEGETSLVVRGRPDGKGL----GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDAT 427
          |:| |    | | |   |          : :  |||||||| |: |    |: :|:
Db    348 CEEFEDGFYVEGLEDISPLKQRFSRIKPPLIKSFNDHRIAMSFAVLTLAL--PLEIDNLE 405

Qy    428 MIATSFPEFMDLM 440
            |||:|  |:
Db    406 CANISFPQFKHLL 418
```

Search completed: August 15, 2002, 14:06:07
Job time: 406 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:56 ; Search time 33.34 Seconds
(without alignments)
1311.357 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1874.5	81.9	480	2	AG3491	3-phosphoshikimate
2	1135	49.6	443	2	C87694	3-phosphoshikimate
3	989.5	43.2	746	2	E83250	still frameshift 3
4	912.5	39.9	454	2	A82572	3-phosphoshikimate
5	884	38.6	431	2	C83858	5-enolpyruvoylshik
6	867.5	37.9	447	2	S76218	3-phosphoshikimate
7	815.5	35.6	428	2	AC1687	5-enolpyruvylshiki
8	806.5	35.2	430	2	H86842	hypothetical prote
9	806	35.2	443	2	S44096	enolpyruvylshikima
10	803.5	35.1	427	2	D95159	3-phosphoshikimate
11	802.5	35.1	430	2	S52580	3-phosphoshikimate
12	801.5	35.0	431	2	D98025	3-phosphoshikimate
13	799.5	34.9	428	2	AC1315	5-enolpyruvylshiki
14	785	34.3	428	2	C26532	5-enolpyruvoylshik
15	766.5	33.5	431	2	D70433	5-enolpyruvylshiki
16	695	30.4	410	2	G72388	3-phosphoshikimate
17	610	26.7	432	2	H89924	3-phosphoshikimate
18	590	25.8	428	2	G81362	3-phosphoshikimate
19	582	25.4	428	2	JC5338	3-phosphoshikimate
20	521.5	22.8	429	2	G71863	3-phosphoshikimate
21	518.5	22.7	429	2	A64570	3-phosphoshikimate
22	496	21.7	419	2	D69202	5-enolpyruvylshiki
23	470.5	20.6	439	2	F84278	3-phosphoshikimate
24	459.5	20.1	436	2	F64362	3-phosphoshikimate
25	408	17.8	462	2	H75438	3-phosphoshikimate
26	403	17.6	416	2	H69436	5-enolpyruvylshiki
27	402.5	17.6	427	1	XUEBY	3-phosphoshikimate
28	402.5	17.6	427	2	A72642	probable 3-phospho
29	401.5	17.5	426	2	D82163	3-phosphoshikimate
30	395.5	17.3	428	2	AH0169	3-phosphoshikimate
31	378	16.5	427	1	XUECVS	3-phosphoshikimate
32	378	16.5	427	2	G90752	5-enolpyruvylshiki
33	378	16.5	427	2	E85616	5-enolpyruvylshiki
34	374.5	16.4	432	2	JN0758	3-phosphoshikimate
35	373.5	16.3	410	2	C75162	3-phosphoshikimate
36	372	16.3	425	2	H97436	3-phosphoshikimate
37	372	16.3	425	2	AD2655	3-phosphoshikimate
38	372	16.3	427	2	S13266	3-phosphoshikimate
39	357.5	15.6	441	2	S28063	3-phosphoshikimate
40	352	15.4	427	1	XUEBVS	3-phosphoshikimate
41	351.5	15.4	428	2	D97010	5-enolpyruvylshiki
42	348	15.2	427	2	A49746	3-phosphoshikimate
43	347	15.2	427	2	AH0613	3-phosphoshikimate
44	342	14.9	427	2	S12096	3-phosphoshikimate
45	337.5	14.8	409	2	S41329	5-enolpyruvylshiki

ALIGNMENTS

RESULT 1
AG3491
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) [imported] - Brucella meli
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AG3491
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov .; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AG3491
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-480 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL53098.1; PID:g17983963; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1917
A;Map position: I
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
C;Keywords: transferase

Query Match 81.9%; Score 1874.5; DB 2; Length 480;
Best Local Similarity 82.1%; Pred. No. 3.2e-112;
Matches 367; Conservative 34; Mismatches 43; Indels 3; Gaps 1;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | |  :||||| |  |:| :||||||||||||||||||:||||||||||||||||
Db      31 MSHSACPKPATARHSQALTGEIRIPGDKSISHRSFMFGGLASGKTRITGLLEGEDVINTG 90

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||||||||| |||:||||| || |||||||||| || |||||||| ||  ::||
Db      91 RAMQAMGARIRKEGDVWIINGVGNGCLLQPEAPLDFGNAGTGARLTMGLVGTYDMKTSFI 150

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||||:: :|||:|:|:|| ||:|  || ||||||||||||||
Db     151 GDASLSKRPMGRVLNPLREMGVQVEAAEGDRMPLTLIGPRTANPIAYRVPMASAQVKSAV 210

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:|||||| ||
Db     211 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRIVGQGKLTGQTID 270

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||:|:||||||||
Db     271 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEIIDPRLAGGED 330

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||::| |||| || :|||||||||:||:||:|||| |||:||:||||||||||:|
Db     331 VADLRVKASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 390
```

```
Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||| |||||   |  | |||||||||||||||| || |
Db     391 VARGLEANGVDCTEGEMSLTVRGRPGGKGLG---GGTVGTHLDHRIAMSFLVMGLASEKP 447

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKI 447
           |||||:|||||||||| :||||||||
Db     448 VTVDDSTMIATSFPEFMGMMAGLGAKI 474


RESULT   2
C87694
3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: C87694
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87694
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-443 <STO>
A;Cross-references: GB:AE005673; NID:g13425333; PIDN:AAK25551.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3589
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox


  Query Match             49.6%;  Score 1135;  DB 2;  Length 443;
  Best Local Similarity   55.8%;  Pred. No. 4.4e-65;
  Matches  251;  Conservative   45;  Mismatches  138;  Indels   16;  Gaps    6;

Qy       2 LHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGK 61
           |  |  | |  |     | | || ||||||||||| : | ||:| | : |||||:||: | :
Db       3 LAGLKSAPGGA-----LRGIVRAPGDKSISHRSMILGALATGTTTVEGLLEGDDVLATAR 57

Qy      62 AMQAMGARIRKEG-DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           |||| |||| :||   | |:  | ||    |   :| ||| || || ||    :   :||
Db      58 AMQAFGARIEREGVGRWRIE--GKGGFEEPVDVIDCGNAGTGVRLIMGAAAGFAMCATFT 115

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           || ||  |||||||:||  ||       |  |||:||:|      : | :||||||||||
Db     116 GDQSLRGRPMGRVLDPLARMGATWLGRDKGRLPLTLKGGNL-RGLNYTLPMASAQVKSAV 174

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRT---IRLEGRGKLTGQ 237
           |||||:  |   |||| ||||||:||: ||| : ||    | :|   :||    ||||
Db     175 LLAGLHAEGGVEVIEPEATRDHTERMLRAFGAEVIVEDRKAGDKTFRHVRLPEGQKLTGT 234

Qy     238 VIDVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAG 297
           : ||||||| |||||||||:||||:||:  |::|  ||||  |||||||| : | |:|
Db     235 HVAVPGDPSSAAFPLVAALIVPGSEVTVEGVMLNELRTGLFTTLQEMGADLVISNVRVAS 294

Qy     298 GEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDR 357
           ||:| |:  | | |||| || :|||||||||||||||||||| | ||| |: |:||||||
Db     295 GEEVGDITARYSQLKGVVVPPERAPSMIDEYPILAVAAAFASGETVMRGVGEMRVKESDR 354

Qy     358 LSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVS 417
           :|  ||||| ||   |     :|     | |     || | || ||||||| |::|: :
Db     355 ISLTANGLKACGVQVVEEPEGFIV----TGTGQPPKGGATVVTHGDHRIAMSHLILGMAA 410

Qy     418 ENPVTVDDATMIATSFPEFMDLMAGLGAKI 447
           :  | ||:  ||||||| | ||| |||| :
Db     411 QAEVAVDEPGMIATSFPGFADLMRGLGATL 440


RESULT   3
E83250
stil frameshift 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE prephenate dehydrogenase F
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83250
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337
A;Accession: E83250
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-746 <STO>
A;Cross-references: GB:AE004740; GB:AE004091; NID:g9949275; PIDN:AAG06552.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA3164


  Query Match             43.2%;  Score 989.5;  DB 2;  Length 746;
  Best Local Similarity   48.5%;  Pred. No. 1.7e-55;
  Matches  214;  Conservative   62;  Mismatches  152;  Indels   13;  Gaps    3;

Qy      12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
           |:    ||||:|:|||||||||| | | || | | : | ||||| : | :| : ||  |
Db     317 AQPGGSLSGTIRVPGDKSISHRSIMLGSLAEGTTEVEGFLEGEDALATIQAFRDMGVVIE 376

Qy      72 -KEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
            :     : |||  || ||  |:  ||: |  ||  ||:    ||||  |||||:||||
Db     377 GPQNGRVTVHGVGLHGLKAPPGPIYLGNSGTSMRLLSGLLAAQPFDSTLTGDASLSKRPM 436

Qy     131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
           ||  ||||||  :::    | |:|:|| :  | : | :|||||||| :|||||   |
Db     437 NRVAKPLREMGAVIETGPEGRPPMTIRGGQRLTGMHYDMPMASAQVKSCLLLAGLYAAGE 496

Qy     191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
           |:| ||  ||||||:||:|||  : ||       | ::|   ||:   |:|| | || ||
Db     497 TSVTEPAPTRDHTERMLRGFGYPVVVEGS-----TAKVESGHKLSATHIEVPADISSAAF 551

Qy     251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
           |||| :  ||:: : :| :|||| |:|  |: || |: : | |  ||| |||:||||:
Db     552 FLVAASIAEGSELVLQHVGINPTRVGVIEILRLMGGDLSLENQREVGGEPVADIRVRSAR 611

Qy     311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
           |||: :|||  |  |||:|:| |||| ||| ||: | ||||||||||:  :|:|||  ||
Db     612 LKGIDIPEDLVPLAIDEFPVLFVAAACAEGRTVLRGAEELRVKESDRIQVMADGLKALGV 671

Qy     371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
           :     :|: |       |   |  |  | |||||||| |  | :  |: : |   :|
Db     672 KAEPTPDGIVIEG-------GAFGGGEVWAHGDHRIAMSFSVASLRASGPIRIHDCANVA 724

Qy     431 TSFPEFMDLMAGLGAKIELSD 451
           |||| |: | |  | :: : :
Db     725 TSFPNFLALCAQTGIRVAVEN 745


RESULT   4
A82572
3-phosphoshikimate 1-carboxyvinyltransferase XF2324 [imported] - Xylella fastidiosa (
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: A82572
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-454 <SIM>
A;Cross-references: GB:AE004043; GB:AE003849; NID:g9107486; PIDN:AAF85123.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
```

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2324
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

Query Match 39.9%; Score 912.5; DB 2; Length 454;
Best Local Similarity 46.9%; Pred. No. 7.1e-51;
Matches 206; Conservative 60; Mismatches 158; Indels 15; Gaps 4;

```
Qy      12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
           | : : | | : |||||||||||: ||  || | :|| | || ||  :| : :  :| ||
Db      23 AHQGTPLHGVLSIPGDKSISHRAVMFAALADGTSRIDGFLEAEDTCSTAEILARLGVRIE 82

Qy      72 KEGDTW-IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
             |  |: |||  || |   ||| ||| || ||  ||:    |||  :||||||:||||
Db      83 TPLSTQRIVHGVGVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVGDASLSKRPM 142

Qy     131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
           || :|| :|| :: : |    |: : | :    | :  |:||||:|||||||||
Db     143 RRVTDPLSQMGARIDTSDDGTPPLRIYGGQLLHGIDFISPVASAQIKSAVLLAGLYARNE 202

Qy     191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
           | | ||  |||:||:||  || :: | |        || |  :|    | :| | || ||
Db     203 TVVREPHPTRDYTERMLTAFGVDIDVSTGC-----ARLRGGQRLCATDITIPADFSSAAF 257

Qy     251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
           |||| ::|||||:|:  | :|| | ||:  |: |||:|  |    ||| | ||||| :
Db     258 YLVAASVIPGSDITLRAVGLNPRRIGLLTVLRLMGANIVESNRHEQGGEPVVDLRVRYAP 317

Qy     311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
           |:|  ||||    ||||:| | |||| ||| ||::|  ||||||||||:|:  ||:: ||
Db     318 LQGTRVPEDLVADMIDEFPALFVAAAAAEGQTVVSGAAELRVKESDRLAAMVTGLRVLGV 377

Qy     371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
             ||      : | | | |  |: |       ||||||:| : | :| : | ::|   :|
Db     378 QVDETADGATIHGGPIGHGTINSHG-------DHRIAMAFSIAGQLSVSTVRIEDVANVA 430

Qy     431 TSFPEFMDL--MAGLGAKI 447
           ||||::  |   || | ::
Db     431 TSFPDYETLARSAGFGLEV 449
```

RESULT 5
C83858
5-enolpyruvoylshikimate-3-phosphate synthase aroE [imported] - Bacillus halodurans (stra
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: C83858
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C83858
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431 <STO>
A;Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB05386.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: aroE
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

Query Match 38.6%; Score 884; DB 2; Length 431;
Best Local Similarity 44.2%; Pred. No. 4.4e-49;
Matches 188; Conservative 72; Mismatches 153; Indels 12; Gaps 5;

```
Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           : || ||:::|||||||||: ||| || | | : | | | | ::|    | :|  | :
Db      10 AKGLKGTIKVPGDKSISHRAVMFGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |  ||  |   || ||: |  || :|::    | |  ||| |: |||| ||
Db      70 ERVTVKGKGWDGLREPSDILDVGNSGTTTRLILGILSTLPFHSVIIGDESIGKRPMKRVT 129

Qy     135 NPLREMGVQVKSED-GDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
            ||: || |:   | |:  |:::|| :    | :  |:||||:|||:|||||   | |:|
Db     130 EPLKSMGAQIDGRDHGNLTPLSIRGGQL-KGIDFHSPVASAQMKSAILLAGLRAEGKTSV 188

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           ||  ||||||:||: || |:    : ||: |: :||   |||| : |||| || || ||
Db     189 TEPAKTRDHTERMLEAFGVNI----EKDGL-TVSIEGGQMLTGQHVVVPGDISSAAFFLV 243

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
           |  :|| | :|: || :|||| |::  |::||| : : | |: ||| |||| : :| |:|
Db     244 AGAMVPHSRITLTNVGINPTRAGILEVLKQMGATLAMENERVQGGEPVADLTIETSVLQG 303

Qy     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373
           | :  |  | :||| ||:|| |  | | ||:   |||:|||::|:  | : |   |
Db     304 VEIGGDIIPRLIDEIPIIAVLATQASGRTVIKDAEELKVKETNRIDTVVSELTKLGASIH 363

Qy     374 EGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSF 433
           :  ::: |    ||     |  |::| |||| |:  :  |::| |||:    || |:
Db     364 ATDDGMIIEGPTPLKG-----GVTVSSHGDHRIGMAMAIAALLAEKPVTVEGTEAIAVSY 418

Qy     434 PEFMD 438
           | | |
Db     419 PSFFD 423
```

RESULT 6
S76218
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Synechocystis sp. (strai
N;Alternate names: 5-enolpyruvylshikimate 3-phosphate synthase
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76218; S33748; S34393
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A;Reference number: S74322; MUID:97061201
A;Accession: S76218
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-447 <KAN>
A;Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BAA18477.1; PID:g165
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
R;Mayes, S.R.; dalla Chiesa, M.; Zhang, Z.; Barber, J.
FEBS Lett. 325, 255-261, 1993
A;Title: The genes aroA and trnQ are located upstream of psbO in the chromosome of Sy
A;Reference number: S33748; MUID:93307506
A;Accession: S33748
A;Molecule type: DNA
A;Residues: 103-194 <MAY>
A;Cross-references: EMBL:X72784; NID:g394747; PIDN:CAA51291.1; PID:g394748
C;Genetics:
A;Gene: aroA
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
C;Keywords: aromatic amino acid biosynthesis; transferase
F;28-439/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

```
  Query Match           37.9%;  Score 867.5;  DB 2;  Length 447;
  Best Local Similarity  48.1%;  Pred. No. 5.2e-48;
  Matches  203;  Conservative   62;  Mismatches  144;  Indels   13;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRK-EGDT 76
           |:| :|:|||||||||: | | :|:||| | ||| ||| :|    :|||| | :   :
Db      26 LTGRLRVPGDKSISHRALMLGAIATGETIIEGLLLGEDPRSTAHCFRAMGAEISELNSEK 85

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLV-GVYDFDSTFIGDASLTKRPMGRVLN 135
           |: | | | |   || ||: |  || :||: |  |    |  || ||  ||| ||:
Db      86 IIVQGRGLGQLQEPSTVLDAGNSGTTMRLMLGLLAGQKDCLFTVTGDDSLRHRPMSRVIQ 145

Qy     136 PLREMGVQV-KSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
           ||::|| ::    :|   |: ::| :   || |  |:|||||| :|||| | | |||
Db     146 PLQQMGAKIWARSNGKFAPLAVQGSQL-KPIHYHSPIASAQVKSCLLLAGLTTEGDTTVT 204

Qy     195 EPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVA 254
           || ::|||:|:||| ||| ||::      :: : |   |||| : |||| || || |||
Db     205 EPALSRDHSERMLQAFGAKLTID---PVTHSVTVHGPAHLTGQRVVVPGDISSAAFWLVA 261

Qy     255 ALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGV 314
           | ::|||:: : || :||||||::  | :|||||   | ||  || |||||||:| |:|
Db     262 ASILPGSELLVENVGINPTRTGVLEVLAQMGADITPENERLVTGEPVADLRVRASHLQGC 321

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
           |   :  | :||| |||||||||||| | :    ||||||||||:|:|: |   |    |
Db     322 TFGGEIIPRLIDEIPILAVAAAFAEGTTRIEDAAELRVKESDRLAAIASELGKMGAKVTE 381

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFP 434
             :  | ::|   |  |    || | :  ||||||:  :  | |     :: |   | |:|
Db     382 FDDGLEIQG---GSPL---QGAEVDSLTDHRIAMALAIAALGSGGQTIINRAEAAAISYP 435

Qy     435 EF 436
           ||
Db     436 EF 437
```

```
RESULT   7
AC1687
5-enolpyruvylshikimate-3-phosphate synthase homolog aroE [imported] - Listeria innocua (
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AC1687
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1687
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97267.1; PID:g16414538; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: aroE
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox
```

```
  Query Match           35.6%;  Score 815.5;  DB 2;  Length 428;
  Best Local Similarity  40.8%;  Pred. No. 1e-44;
  Matches  173;  Conservative   83;  Mismatches  155;  Indels   13;  Gaps    5;

Qy      17 GLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT 76
           || | : :|||||:|||| ||| :| |:| |   |  :| : | || :|:| :| :  :
Db       9 GLVGEITVPGDKSMSHRSIMFGAIAEGKTVIRHFLRADDCLGTIKAFKALGVKIEETDEE 68

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNP 136
           |: | |: ||   | ||| ||: |  || ||::   |||:  :|| |: |||| ||: |
```

```
Db      69 IIVHGTGSDGLKQAEGPLDIGNSGTTIRLMMGILAGRDFDTVILGDESIAKRPMNRVMLP 128

Qy     137 LREMGVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
           |:||| ::  :|    |::: | ::   :  | :|:|||||||:: | |   | | : |
Db     129 LQEMGAKMHGKDGSEFAPISIIGNQSLKRMEYHMPVASAQVKSAIIFAALQAEGETIIHE 188

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
           ||||||  |:: ||     | : ||: |||::|  |  || : |||| || || :||
Db     189 KEKTRDHTEHMIRQFGG----EIEMDGL-TIRVKGGQKFIGQEMTVPGDVSSAAFFIVAG 243

Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
           |: |||:: : :| :||||||:   :::||  : | :   : |:    : |:|| |||
Db     244 LITPGSEIELTHVGLNPTRTGIFDVVEQMGGSLVVKDSSRSTGKLAGTVVVKSSELKGTE 303

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
           :  |  | :||| |::|: |  ||| |::    ||:|||::|: |||| |   | |
Db     304 IGGDIIPRLIDEIPVIALLATQAEGTTIIKDAAELKVKETNRIDAVANELNKMGADITPT 363

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSFP 434
           |  |::||:           | | :: |||| |   :   || :  | :| |  :: |:|
Db     364 EDGLIIRGKTP------LHAANVTSYGDHRIGMMLQIAALLVEDGDVELDRAEAVSVSYP 417

Qy     435 EFMD 438
           | :
Db     418 TFFE 421
```

```
RESULT   8
H86842
hypothetical protein aroA [imported] - Lactococcus lactis subsp. lactis (strain IL140
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: H86842
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86842
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <STO>
A;Cross-references: GB:AE005176; PID:g12724764; PIDN:AAK05842.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: aroA
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
```

```
  Query Match           35.2%;  Score 806.5;  DB 2;  Length 430;
  Best Local Similarity  41.9%;  Pred. No. 3.8e-44;
  Matches  184;  Conservative   77;  Mismatches  157;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| |||||::| :| :|||  |  :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTVIYDILRGEDVLSTIEAFRAMGVEIEDKG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |   | |||  || ||: |  ||  |::    |::|  || ||:|||| ||
Db      67 EVITVHGKGISELKAPEKALDMGNSGTSTRLLSGILAGLPFETTLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|       | | :|:|||||||||: | |   |:| |
Db     127 TPLQLMGAEITGQTDKVKLPMTIKGSTHLKAIDYVLPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           :|   || |||:||  ||  | |       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGELKVSD-----KTILVPGGQKLVGQKVVVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           |||:|  |:: : || :| ||||:|  :|  ||  :|::       :|     | |:|::|
Db     242 AALVVENSELILENVGVNETRTGIIEVIQAMGGQLEILE------QDNVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
```

```
Db    296 QLKGTEISGDLIPRLIDELPIIALLATQAQGQTIIRDAAELKVKETDRIAVVANALNSMG 355

Qy    370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428

Db    356 AKIQPTDDGMIIQG---GTKL-HAPENSINTLGDHRIGMMAAIAALLVKNGEIELERAEA 411

Qy    429 IATSFPEFMDLMAGLGAKI 447

Db    412 IQTSYPSFFDDLEQLSENI 430


RESULT   9
S44096
enolpyruvylshikimate 3-phosphate synthase - Dichelobacter nodosus
C;Species: Dichelobacter nodosus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C;Accession: S44096
R;Alm, R.A.; Dalrymple, B.P.; Mattick, J.S.
Submitted to the EMBL Data Library, January 1994
A;Description: Sequencing and expression of the aroA gene from Dichelobacter nodosus.
A;Reference number: S44096
A;Accession: S44096
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-443 <ALM>
A;Cross-references: EMBL:Z29339; NID:g472931; PIDN:CAA82544.1; PID:g472932
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox
F;16-424/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>


  Query Match             35.2%;  Score 806;  DB 2;  Length 443;
  Best Local Similarity   42.2%;  Pred. No. 4.3e-44;
  Matches  185;  Conservative   74;  Mismatches  169;  Indels   10;  Gaps    3;

Qy     16 SGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75

Db     12 SALSGEITICGDKSMSHRALLLAALAEGQTEIRGFLACADCLATRQALRALGVDIQREKE 71

Qy     76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135

Db     72 IVTIRGVGFLGLQPPKAPLNMQNSGTSMRLLAGILAAQRFESVLCGDESLEKRPMQRIIT 131

Qy    136 PLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195

Db    132 PLVQMGAKIVSHSNFTAPLHISG-RPLTGIDYALPLPSAQLKSCLILAGLLADGTTRLHT 190

Qy    196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255

Db    191 CGISRDHTERMLPLFGGALEIKKE-----QIIVTGGQKLHGCVLDIVGDLSAAAFFMVAA 245

Qy    256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315

Db    246 LIAPRAEVVIRNVGINPTRAAIITLLQKMGGRIELHHQRFWGAEPVADIVVYHSKLRGIT 305

Qy    316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375

Db    306 VAPEWIANAIDELPIFFIAAACAEGTTFVGNLSELRVKESDRLAAMAQNLQTLGVACDVG 365

Qy    376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPE 435

Db    366 ADFIHIYGRSDRQFL----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMPQ 421

Qy    436 FMDLMAGLGAKIELSDTK 453

Db    422 FRDFAAAIGMNVGEKDAK 439


RESULT  10
D95159
3-phosphoshikimate 1-carboxyvinyltransferase [imported] - Streptococcus pneumoniae (stra
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: D95159
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: D95159
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75469.1; PID:g14972856; GSPDB:GN00164; TIGR:
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1371
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car


  Query Match             35.1%;  Score 803.5;  DB 2;  Length 427;
  Best Local Similarity   42.8%;  Pred. No. 5.9e-44;
  Matches  184;  Conservative   72;  Mismatches  155;  Indels   19;  Gaps    7;

Qy     18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77

Db     10 LHGSIRVPGDKSISHRSIIFGSLAEGETKVYDILRGEDVLSTMQVFRDLGVEIEDKDGVI 69

Qy     78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137

Db     70 TIQGVGMAGLKAPQNALNMGNSGTSIRLISGVLAGADFEVEMFGDDSLSKRPMDRVTLPL 129

Qy    138 REMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196

Db    130 KKMGVSISGQTERDLPPLRLKGTKNLRPIHYELPIASAQVKSALMFAALQAKGESVIIEK 189

Qy    197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256

Db    190 EYTRNHTEDMLKQFGGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISSAAFWLVAGL 244

Qy    257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKGV 314

Db    245 IAPNSRLVLQNVGINETRTGIIDVIRAMGGKLEITEIDPVAKS----ATLIVESSDLKGT 300

Qy    315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374

Db    301 EIGGALIPRLIDELPIIALLATQAQGVTVIKDAEELKVKETDRIQVVADALNSMGADITP 360

Qy    375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSF 433

Db    361 TADGMIIKGK------SALHGARVNTFGDHRIGMMTAIAALLVADGEVELDRAEAINTSY 414

Qy    434 PEFMDLMAGL 443

Db    415 PSFFDDLESL 424


RESULT  11
S52580
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Lactococcus lactis
N;Alternate names: 5-enolpyruvylshikimate-3-phosphate synthase
C;Species: Lactococcus lactis
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C;Accession: S52580
R;Griffin, H.G.; Gasson, M.J.
Mol. Gen. Genet. 246, 119-127, 1995
A;Title: Genetic aspects of aromatic amino acid biosynthesis in Lactococcus lactis.
A;Reference number: S52579; MUID:95124293
A;Accession: S52580
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <GRI>
A;Cross-references: EMBL:X78413; NID:g683581; PIDN:CAA55180.1; PID:g683583
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
```

```
C;Keywords: transferase
F;12-421/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

  Query Match           35.1%;  Score 802.5;  DB 2;  Length 430;
  Best Local Similarity  41.2%;  Pred. No. 6.9e-44;
  Matches  181;  Conservative   79;  Mismatches  158;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| ||||::| :| :|:|  |  :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTIIHDILRGEDVLSTIEAFRALGVEIEDDG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
             : | |   |  ||  || ||: |  ||  |::    |::|  || ||:|||| ||
Db      67 QVITVHGQGISKLKEPEKALDMGNSGTSTRLLSGILAGLPFEATLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|      | | :|:|||||||||: | |   |:| |
Db     127 TPLQMMGAEIVGQTDKVKLPMTIKGSAHLKAIDYILPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           :|   || |||:||  ||  :||       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGEITVSD-----KTILVPGGQKLLGQEVTVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           | |:|  | : : || :| ||||::  :| ||  :|::       :|     | |:|::|
Db     242 AGLVVENSGLILENVGINETRTGILEVIQAMGGQLEILE------QDEVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
           |||  :  |  | :||| ||:|: |  ||| |::    ||:|||:||:: ||: |  |
Db     296 QLKGTEISGDLIPRLIDELPIIALLATQAEGKTIIRDAAELKVKETDRIAVVADALNSMG 355

Qy     370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428
           : :  :  ::::|   | | :|  |: |  |||| |   :  |: ||  : :: |
Db     356 ANIEPTDDGMIIQG---GTKL-HAPENAINTLGDHRIGMMVAIAALLVENGEIELERAEA 411

Qy     429 IATSFPEFMDLMAGLGAKI 447
           | ||:| | | :  |   :
Db     412 IQTSYPSFFDDLEKLSGNL 430


RESULT  12
D98025
3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) [imported] - Streptococcus pn
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: D98025
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
  R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
  P.; Sun, P.M.; Winkler, M.E.
  J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D98025
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-431 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00033.1; PID:g15458865; GSPDB:GN00174
C;Genetics:
A;Gene: aroA
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox
C;Keywords: transferase

  Query Match           35.0%;  Score 801.5;  DB 2;  Length 431;
  Best Local Similarity  42.8%;  Pred. No. 8e-44;
  Matches  184;  Conservative   72;  Mismatches  155;  Indels   19;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
           | |::):|||||||||| :|| || |||::  :| ||||::| :  : :|  |  :
Db      14 LHGSIRVPGDKSISHRSIIFGSLAEGETKVYDILRGEDVLSTMQVFRDLGVEIEDKDGVI 73

Qy      78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
           | |||  || ||:  |: ||: |  ||  |::   ||:    || ||:|||| ||  ||
Db      74 TIQGVGMAGLKAPQNALNMGNSGTSIRLISGVLAGADFEVEMFGDDSLSKRPMDRVTLPL 133

Qy     138 REMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196
           ::||| :  : : |  |: |:| |   || | :|:|||||||:: | |   | : :||
Db     134 KKMGVSISGQTERDLPPLRLKGTKNLRPIHYELPIASAQVKSALMFAALQAKGESVIIEK 193

Qy     197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
             ||:||| ||: || :|:|    || : | ::|  ||||| : |||| || || ||| |
Db     194 ECTRNHTEDMLKQFGGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISSAAFWLVAGL 248

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKGV 314
           : | | : : || :| ||||:|  :: ||  :||  |:|         : | | || |||
Db     249 INPNSHLVLQNVGINETRTGIIDVIRAMGGKLEVTEIDPVAKS----STLTVESSDLKGT 304

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
           :    | :||| ||:|: |  |:| ||:  |||:|||:||:  ||: |  | |
Db     305 EIGGALIPRLIDELPIIALLATQAQGVTVIKDAEELKVKETDRIQVVADALNSMGADITP 364

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSF 433
           :::::|:                || | |  |||| |   :   ||::  | :| |  | ||:
Db     365 TADGMIIKGK------SALHGARVNTFGDHRIGMMTAIAALLVADGEVELDRAEAINTSY 418

Qy     434 PEFMDLMAGL 443
           | | | :  |
Db     419 PSFFDDLESL 428


RESULT  13
AC1315
5-enolpyruvylshikimate-3-phosphate synthase homolog aroE [imported] - Listeria monocy
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AC1315
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1315
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-428 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00001.1; PID:g16411376; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: aroE
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car

  Query Match           34.9%;  Score 799.5;  DB 2;  Length 428;
  Best Local Similarity  39.9%;  Pred. No. 1.1e-43;
  Matches  169;  Conservative   84;  Mismatches  158;  Indels   13;  Gaps    5;

Qy      17 GLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT 76
           || | : :|||||:|||| ||| :| |:| |   |  :| : | || :|:| :| :  :
Db       9 GLVGAITVPGDKSMSHRSIMFGAIAEGKTVIRHFLRADDCLGTIKAFKALGVKIEETDEE 68

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNP 136
             |: | |  ||   : ||| ||: |  || ||::   |||:  :|| |: |||| ||: |
Db      69 IIVHGTGFDGLKQADGPLDIGNSGTTIRLMMGILAGRDFDTVILGDESIAKRPMNRVMLP 128

Qy     137 LREMGVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
           |::|| ::  :||    |:|: | ::   : | :|:|||||||||:: | |   | : |
Db     129 LQQMGAKMHGKDGSEFAPITITGKQSLKRMEYHMPVASAQVKSAIIFAALQAEGETIIHE 188

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
             |||||| |:: ||     | : ||: |||::|  | ||| : |||| || || :||
Db     189 KEKTRDHTEHMIRQFGG----EIEMDGL-TIRVKGGQKFTGQEMTVPGDVSSAAFFIVAG 243
```

```
Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
           |: |||:: : :| :|||||:|:   :::|| : | :   : |:    : |:::| |||
Db     244 LITPGSEIELTHVGLNPTRTGIFDVVKQMGGSLVVKDSSRSTGKLAGTVVVKTSKLKGTE 303

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
           :  |  | :||| |::|: |  ||| |::    ||:|||::|: |||  |   | |
Db     304 IDGDIIPRLIDEIPVIALLATQAEGTTIIKDAAELKVKETNRIDAVATELNKMGADITPT 363

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSFP 434
           |  |:: |:           | | :: |||| |   :   || | | ::    :: |:|
Db     364 EDGLIIHGKTP------LHAANVTSYGDHRIGMMLQIAALLVEEGDVELERPEAVSVSYP 417

Qy     435 EFMD 438
             | :
Db     418 TFFE 421


RESULT  14
C26532
5-enolpyruvoylshikimate-3-phosphate synthase aroE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 03-Nov-1987 #sequence_revision 03-Nov-1987 #text_change 20-Jun-2000
C;Accession: C26532; B69590
R;Henner, D.J.; Band, L.; Flaggs, G.; Chen, E.
Gene 49, 147-152, 1986
A;Title: The organization and nucleotide sequence of the Bacillus subtilis hisH, tyrA an
A;Reference number: A26532; MUID:87191994
A;Accession: C26532
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-428 <HEN>
A;Cross-references: GB:M80245; GB:M15409; NID:g143798; PIDN:AAA20869.1; PID:g143816
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033
A;Accession: B69590
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-428 <KUN>
A;Cross-references: GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB14176.1; PID:g2634678
A;Experimental source: strain 168
C;Genetics:
A;Gene: aroE
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox
F;11-419/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>


  Query Match             34.3%;  Score 785;  DB 2;  Length 428;
  Best Local Similarity   41.0%;  Pred. No. 9e-43;
  Matches  178;  Conservative   68;  Mismatches  176;  Indels   12;  Gaps    4;

Qy      14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73
           |   | | : |||||||||||| ||| ||:| | :   | | | ::|    : ||  | :
Db       5 KVQTLHGEIHIPGDKSISHRSVMFGALAAGTTTVKNFLPGADCLSTIDCFRKMGVHIEQS 64

Qy      74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133
              :| | |   |  ||: || ||: |  || :|::    | |   || |: |||| ||
Db      65 SSDVVIHGKGIDALKEPESLLDVGNSGTTIRLMLGILAGRPFYSAVAGDESIAKRPMKRV 124

Qy     134 LNPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192
             ||::|| ::      |:  |::: |       | |  |:||||:|||||||||   | ||
Db     125 TEPLKKMGAKIDGRAGGEFTPLSVSGASL-KGIDYVSPVASAQIKSAVLLAGLQAEGTTT 183

Qy     193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252
           | ||  :|||||:||  ||  |     ::   :: : |  |||   | ||||  || || |
Db     184 VTEPHKSRDHTERMLSAFGVKL-----SEDQTSVSIAGGQKLTAADIFVPGDISSAAFFL 238

Qy     253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLK 312
           |  :|| | : :  || :|||||||:|  || ||| :|:     :| |   || : :|:||
Db     239 AAGAMVPNSRIVLKNVGLNPTRTGIIDVLQNMGAKLEIKPSADSGAEPYGDLIIETSSLK 298

Qy     313 GVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDC 372
           | :  |  | :||| ||:|: |  ||| ||:    ||:|||::|:  | : |:  | :
Db     299 AVEIGGDIIPRLIDEIPIIALLATQAEGTTVIKDAAELKVKETNRIDTVVSELRKLGAEI 358

Qy     373 DEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATS 432
           :     : | |:   ||     ||||::| |||| |   :   ::| |: ::    |  |
Db     359 EPTADGMKVYGKQTLKG-----GAAVSSHGDHRIGMMLGIASCITEEPIEIEHTDAIHVS 413

Qy     433 FPEFMDLMAGLGAK 446
           :| | : :  |  |
Db     414 YPTFFEHLNKLSKK 427


RESULT  15
D70433
5-enolpyruvylshikimate-3-phosphate synthetase - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 18-Jun-1999
C;Accession: D70433
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666
A;Accession: D70433
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-431 <AQF>
A;Cross-references: GB:AE000744; NID:g2983891; PIDN:AAC07443.1; PID:g2983892; GB:AE00
A;Experimental source: strain VF5
C;Genetics:
A;Gene: aroA
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-car
F;12-423/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>


  Query Match             33.5%;  Score 766.5;  DB 2;  Length 431;
  Best Local Similarity   40.7%;  Pred. No. 1.4e-41;
  Matches  174;  Conservative   73;  Mismatches  169;  Indels   11;  Gaps    6;

Qy      14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73
           |   : | :|:| ||||:||:|: | |||||| :   |   | : | : ::|:  ::|:
Db       6 KIKRVKGELRVPSDKSITHRAFILGALASGETLVRKPLISGDTLATLEILKAIRTKVREG 65

Qy      74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133
           :  :|:|  |   | |   ||  |: |  |:  |::    | |   || || ||| ||
Db      66 KEEVLIEG-RNYTFLEPHDVLDAKNSGTTARIMSGVLSTQPFFSVLTGDESLKNRPMLRV 124

Qy     134 LNPLREMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192
           : |||||| ::   |:|::||: :||       |:|    :|||||||:|||||   |:|
Db     125 VEPLREMGAKIDGREEGNKLPIAIRGGNL-KGISYFNKKSSAQVKSALLLAGLRAEGMTE 183

Qy     193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252
           |:|| ::|||||:||: ||| :    :  |   ::::|  :| |  :  | |||| |:
Db     184 VVEPYLSRDHTERMLKLFGAEVITIPEERG-HIVKIKGGQELQGTEVYCPADPSSAAYFA 242

Qy     253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRS-STL 311
           | | |  :: :  ||:|||| |    | ||| ||   | |    | :||| ||    |
Db     243 ALATLAPEGEIRLKEVLLNPTRDGFYRKLIEMGGDISFENYRELSNEPMADLVVRPVDNL 302

Qy     312 KGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVD 371
```

```
          | | |  :  |::||| |||||  |||:| : : | :||| |||||: |:   |:  ||
Db    303 KPVKVSPEEVPTLIDEIPILAVLMAFADGVSEVKGAKELRYKESDRIKAIVTNLRKLGVQ 362

Qy    372 CDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIAT 431
          :| |    : |  : |      |  : |  ||||||:| |:||| |  | :|    :
Db    363 VEEFEDGFAIHGTKEIK------GGVIETFKDHRIAMAFAVLGLVVEEEVIIDHPECVTV 416

Qy    432 SFPEFMD 438
          |:||| :
Db    417 SYPEFWE 423


Search completed: August 15, 2002, 13:59:57
Job time: 131 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:16 ; Search time 22.39 Seconds
(without alignments)
496.366 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

al number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2282	99.7	455	1	US-08-476-008-3	Sequence 3, Appli
2	2282	99.7	455	1	US-08-306-063-3	Sequence 3, Appli
3	2282	99.7	455	1	US-08-833-485-3	Sequence 3, Appli
4	2282	99.7	455	4	US-09-137-440-3	Sequence 3, Appli
5	2282	99.7	455	5	PCT-US91-06148A-3	Sequence 3, Appli
6	1900.5	83.1	449	1	US-08-476-008-5	Sequence 5, Appli
7	1900.5	83.1	449	1	US-08-476-008-7	Sequence 7, Appli
8	1900.5	83.1	449	1	US-08-306-063-5	Sequence 5, Appli
9	1900.5	83.1	449	1	US-08-306-063-7	Sequence 7, Appli
10	1900.5	83.1	449	1	US-08-833-485-5	Sequence 5, Appli
11	1900.5	83.1	449	1	US-08-833-485-7	Sequence 7, Appli
12	1900.5	83.1	449	4	US-09-137-440-5	Sequence 5, Appli
13	1900.5	83.1	449	4	US-09-137-440-7	Sequence 7, Appli
14	1900.5	83.1	449	5	PCT-US91-06148A-5	Sequence 5, Appli
15	1900.5	83.1	449	5	PCT-US91-06148A-7	Sequence 7, Appli
16	867.5	37.9	447	1	US-08-476-008-67	Sequence 67, Appl
17	867.5	37.9	447	1	US-08-306-063-67	Sequence 67, Appl
18	867.5	37.9	447	1	US-08-833-485-67	Sequence 67, Appl
19	867.5	37.9	447	4	US-09-137-440-67	Sequence 67, Appl
20	806	35.2	443	1	US-08-476-008-69	Sequence 69, Appl
21	806	35.2	443	1	US-08-306-063-69	Sequence 69, Appl
22	806	35.2	443	1	US-08-833-485-69	Sequence 69, Appl
23	806	35.2	443	4	US-09-137-440-69	Sequence 69, Appl
24	803.5	35.1	427	4	US-09-325-881-2	Sequence 2, Appli
25	798.5	34.9	427	2	US-08-896-345-2	Sequence 2, Appli
26	798.5	34.9	427	4	US-09-226-091-2	Sequence 2, Appli
27	785	34.3	428	1	US-08-476-008-42	Sequence 42, Appl
28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appli
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appli
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appli
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appli
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

ALIGNMENTS

```
RESULT 1
US-08-476-008-3
; Sequence 3, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,008
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-08-476-008-3


  Query Match             99.7%;  Score 2282;  DB 1;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-202;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   2
US-08-306-063-3
; Sequence 3, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-3


  Query Match             99.7%;  Score 2282;  DB 1;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-202;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   3
US-08-833-485-3
; Sequence 3, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
```

```
;       STATE:  Missouri
;       COUNTRY:  USA
;       ZIP:  63198
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/833,485
;       FILING DATE:  07-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/306,063
;       FILING DATE:  13-SEP-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/749,611
;       FILING DATE:  28-AUG-1991
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/576,537
;       FILING DATE:  31-AUG-1990
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hoerner Jr., Dennis R.
;       REGISTRATION NUMBER:  30,914
;       REFERENCE/DOCKET NUMBER:  38-21(15117)A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314)737-6099
;       TELEFAX:  (314)737-6047
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  455 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-833-485-3


  Query Match             99.7%;  Score 2282;  DB 1;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-202;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   4
US-09-137-440-3
; Sequence 3, Application US/09137440
; Patent No. 6248876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-3


  Query Match             99.7%;  Score 2282;  DB 4;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.3e-202;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||||
Db    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   5
PCT-US91-06148A-3
; Sequence 3, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06148A
;      FILING DATE:  19910828
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10535)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-06148A-3
```

```
  Query Match            99.7%;  Score 2282;  DB 5;  Length 455;
  Best Local Similarity  99.8%;  Pred. No. 1.3e-202;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||||
Db    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   6
US-08-476-008-5
; Sequence 5, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,008
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-008-5

  Query Match             83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   7
US-08-476-008-7
; Sequence 7, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,008
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-008-7

  Query Match             83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   8
US-08-306-063-5
```

```
; Sequence 5, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-5


  Query Match           83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity  82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||    | |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   9
US-08-306-063-7
; Sequence 7, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-7


  Query Match           83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity  82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
```

```
Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: |||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  10
US-08-833-485-5
; Sequence 5, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,485
;      FILING DATE:  07-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-833-485-5


  Query Match             83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  11
US-08-833-485-7
; Sequence 7, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/833,485
;       FILING DATE:  07-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/306,063
;       FILING DATE:  13-SEP-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/749,611
;       FILING DATE:  28-AUG-1991
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/576,537
;       FILING DATE:  31-AUG-1990
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hoerner Jr., Dennis R.
;       REGISTRATION NUMBER:  30,914
;       REFERENCE/DOCKET NUMBER:  38-21(15117)A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314)737-6099
;       TELEFAX:  (314)737-6047
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  449 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-833-485-7


  Query Match             83.1%;  Score 1900.5;  DB 1;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   |  |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  12
US-09-137-440-5
; Sequence 5, Application US/09137440
```

```
; Patent No. 6248876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-5


  Query Match             83.1%;  Score 1900.5;  DB 4;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
```

```
          |||||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db    241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :|||||||||||:||:||:|||| |||:||:|||||||||||:|
Db    301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db    361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy    421 VTVDDATMIATSFPEEFMDLMAGLGAKIELS 450
          |||||: ||||||||||:| |||||||||
Db    418 VTVDDSNMIATSFPEEFMDMMPGLGAKIELS 447


RESULT  13
US-09-137-440-7
; Sequence 7, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-7


  Query Match             83.1%;  Score 1900.5;  DB 4;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          | | || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||||
Db      1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          :|||||||:|||||| |||:|||||| || ||| |||||| || |||||||| ||  ::||
Db     61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||:||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db    121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:|||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db    241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :|||||||||||:||:||:|||| |||:||:|||||||||||:|
Db    301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db    361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy    421 VTVDDATMIATSFPEEFMDLMAGLGAKIELS 450
          |||||: ||||||||||:| |||||||||
Db    418 VTVDDSNMIATSFPEEFMDMMPGLGAKIELS 447


RESULT  14
PCT-US91-06148A-5
; Sequence 5, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06148A
;      FILING DATE:  19910828
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10535)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
```

```
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-06148A-5


  Query Match             83.1%;  Score 1900.5;  DB 5;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: |||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   |  ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  15
PCT-US91-06148A-7
; Sequence 7, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06148A
;      FILING DATE:  19910828
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10535)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-06148A-7


  Query Match             83.1%;  Score 1900.5;  DB 5;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 2.2e-167;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: |||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   |  ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


Search completed: August 15, 2002, 13:59:17
Job time: 91 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:46 ; Search time 22.39 Seconds
(without alignments)
496.366 Million cell updates/sec

Title: US-09-464-099A-3
Perfect score: 2288
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2288	100.0	455	1	US-08-476-008-3	Sequence 3, Appli
2	2288	100.0	455	1	US-08-306-063-3	Sequence 3, Appli
3	2288	100.0	455	1	US-08-833-485-3	Sequence 3, Appli
4	2288	100.0	455	4	US-09-137-440-3	Sequence 3, Appli
5	2288	100.0	455	5	PCT-US91-06148A-3	Sequence 3, Appli
6	1906.5	83.3	449	1	US-08-476-008-5	Sequence 5, Appli
7	1906.5	83.3	449	1	US-08-476-008-7	Sequence 7, Appli
8	1906.5	83.3	449	1	US-08-306-063-5	Sequence 5, Appli
9	1906.5	83.3	449	1	US-08-306-063-7	Sequence 7, Appli
10	1906.5	83.3	449	1	US-08-833-485-5	Sequence 5, Appli
11	1906.5	83.3	449	1	US-08-833-485-7	Sequence 7, Appli
12	1906.5	83.3	449	4	US-09-137-440-5	Sequence 5, Appli
13	1906.5	83.3	449	4	US-09-137-440-7	Sequence 7, Appli
14	1906.5	83.3	449	5	PCT-US91-06148A-5	Sequence 5, Appli
15	1906.5	83.3	449	5	PCT-US91-06148A-7	Sequence 7, Appli
16	867.5	37.9	447	1	US-08-476-008-67	Sequence 67, Appl
17	867.5	37.9	447	1	US-08-306-063-67	Sequence 67, Appl
18	867.5	37.9	447	1	US-08-833-485-67	Sequence 67, Appl
19	867.5	37.9	447	4	US-09-137-440-67	Sequence 67, Appl
20	806	35.2	443	1	US-08-476-008-69	Sequence 69, Appl
21	806	35.2	443	1	US-08-306-063-69	Sequence 69, Appl
22	806	35.2	443	1	US-08-833-485-69	Sequence 69, Appl
23	806	35.2	443	4	US-09-137-440-69	Sequence 69, Appl
24	803.5	35.1	427	4	US-09-325-881-2	Sequence 2, Appli
25	798.5	34.9	427	2	US-08-896-345-2	Sequence 2, Appli
26	798.5	34.9	427	4	US-09-226-091-2	Sequence 2, Appli
27	785	34.3	428	1	US-08-476-008-42	Sequence 42, Appl
28	785	34.3	428	1	US-08-306-063-42	Sequence 42, Appl
29	785	34.3	428	1	US-08-833-485-42	Sequence 42, Appl
30	785	34.3	428	4	US-09-137-440-42	Sequence 42, Appl
31	704.5	30.8	415	2	US-08-896-345-4	Sequence 4, Appli
32	704.5	30.8	415	4	US-09-226-091-4	Sequence 4, Appli
33	704.5	30.8	415	4	US-09-325-881-4	Sequence 4, Appli
34	577	25.2	430	1	US-08-476-008-44	Sequence 44, Appl
35	577	25.2	430	1	US-08-306-063-44	Sequence 44, Appl
36	577	25.2	430	1	US-08-833-485-44	Sequence 44, Appl
37	577	25.2	430	4	US-09-137-440-44	Sequence 44, Appl
38	402.5	17.6	427	1	US-08-476-008-60	Sequence 60, Appl
39	402.5	17.6	427	1	US-08-306-063-60	Sequence 60, Appl
40	402.5	17.6	427	1	US-08-833-485-60	Sequence 60, Appl
41	402.5	17.6	427	4	US-09-137-440-60	Sequence 60, Appl
42	378	16.5	427	4	US-09-243-374-7	Sequence 7, Appli
43	374.5	16.4	432	1	US-08-476-008-61	Sequence 61, Appl
44	374.5	16.4	432	1	US-08-306-063-61	Sequence 61, Appl
45	374.5	16.4	432	1	US-08-833-485-61	Sequence 61, Appl

ALIGNMENTS

```
RESULT   1
US-08-476-008-3
; Sequence 3, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,008
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
US-08-476-008-3


  Query Match             100.0%;  Score 2288;  DB 1;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-203;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   2
US-08-306-063-3
; Sequence 3, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-3


  Query Match             100.0%;  Score 2288;  DB 1;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-203;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   3
US-08-833-485-3
; Sequence 3, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
```

```
;       STATE:  Missouri
;       COUNTRY:  USA
;       ZIP:  63198
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/833,485
;       FILING DATE:  07-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/306,063
;       FILING DATE:  13-SEP-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/749,611
;       FILING DATE:  28-AUG-1991
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/576,537
;       FILING DATE:  31-AUG-1990
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hoerner Jr., Dennis R.
;       REGISTRATION NUMBER:  30,914
;       REFERENCE/DOCKET NUMBER:  38-21(15117)A
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314)737-6099
;       TELEFAX:  (314)737-6047
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  455 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-833-485-3


  Query Match           100.0%;  Score 2288;  DB 1;  Length 455;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-203;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455


RESULT   4
US-09-137-440-3
; Sequence 3, Application US/09137440
; Patent No. 6248876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  455 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-3


  Query Match           100.0%;  Score 2288;  DB 4;  Length 455;
  Best Local Similarity  100.0%;  Pred. No. 2.6e-203;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||||
Db    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   5
PCT-US91-06148A-3
; Sequence 3, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;     APPLICANT:  Barry, Gerard F.
;     APPLICANT:  Kishore, Ganesh M.
;     APPLICANT:  Padgette, Stephen R.
;     TITLE OF INVENTION:  Glyphosate Tolerant
;     TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;     NUMBER OF SEQUENCES:  36
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;       STREET:  700 Chesterfield Village Parkway
;       CITY:  St. Louis
;       STATE:  Missouri
;       COUNTRY:  USA
;       ZIP:  63198
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US91/06148A
;       FILING DATE:  19910828
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/749,611
;       FILING DATE:  28-AUG-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/576,537
;       FILING DATE:  31-AUG-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hoerner Jr., Dennis R.
;       REGISTRATION NUMBER:  30,914
;       REFERENCE/DOCKET NUMBER:  38-21(10535)
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314)537-6099
;       TELEFAX:  (314)537-6047
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  455 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
PCT-US91-06148A-3


  Query Match             100.0%;  Score 2288;  DB 5;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-203;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||||
Db    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   6
US-08-476-008-5
; Sequence 5, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;     APPLICANT:  Barry, Gerard F.
;     APPLICANT:  Kishore, Ganesh M.
;     APPLICANT:  Padgette, Stephen R.
;     APPLICANT:  Stallings, William C.
;     TITLE OF INVENTION:  Glyphosate Tolerant
;     TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;     NUMBER OF SEQUENCES:  69
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;       STREET:  700 Chesterfield Village Parkway
;       CITY:  St. Louis
;       STATE:  Missouri
;       COUNTRY:  USA
;       ZIP:  63198
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/476,008
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/306,063
;       FILING DATE:  13-SEP-1994
;       APPLICATION NUMBER:  US 07/749,611
;       FILING DATE:  28-AUG-1991
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/576,537
;       FILING DATE:  31-AUG-1990
;       CLASSIFICATION:  435
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-008-5


  Query Match             83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   7
US-08-476-008-7
; Sequence 7, Application US/08476008
; Patent No. 5627061
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/476,008
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-476-008-7


  Query Match             83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   8
US-08-306-063-5
```

```
; Sequence 5, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-5


  Query Match           83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity  83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   |  |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT   9
US-08-306-063-7
; Sequence 7, Application US/08306063
; Patent No. 5633435
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10660)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-306-063-7


  Query Match           83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity  83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
```

```
Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  10
US-08-833-485-5
; Sequence 5, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,485
;      FILING DATE:  07-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-833-485-5


  Query Match             83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  11
US-08-833-485-7
; Sequence 7, Application US/08833485
; Patent No. 5804425
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/833,485
;      FILING DATE:  07-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-833-485-7


  Query Match             83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||    |  |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  12
US-09-137-440-5
; Sequence 5, Application US/09137440
; Patent No. 6248876
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-5


  Query Match             83.3%;  Score 1906.5;  DB 4;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
```

```
          |||||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db    241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db    301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||    |  ||||||||||||||||||| :| |
Db    361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: ||||||||||:| |||||||||
Db    418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  13
US-09-137-440-7
; Sequence 7, Application US/09137440
; Patent No. 6248876
; GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    APPLICANT:  Stallings, William C.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/137,440
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/833,485
;      FILING DATE:  07-APR-1997
;      APPLICATION NUMBER:  US 08/306,063
;      FILING DATE:  13-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(15117)A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)737-6099
;      TELEFAX:  (314)737-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-137-440-7


  Query Match             83.3%;  Score 1906.5;  DB 4;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||| || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||||
Db      1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          :|||||||:|||||| |||:|||| || ||| |||||| || |||||||| ||  ::||
Db     61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||:|||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db    121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db    241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db    301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||    |  ||||||||||||||||||| :| |
Db    361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: ||||||||||:| |||||||||
Db    418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  14
PCT-US91-06148A-5
; Sequence 5, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06148A
;      FILING DATE:  19910828
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10535)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
```

```
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-06148A-5


  Query Match             83.3%;  Score 1906.5;  DB 5;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:|||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


RESULT  15
PCT-US91-06148A-7
; Sequence 7, Application PC/TUS9106148A
;  GENERAL INFORMATION:
;    APPLICANT:  Barry, Gerard F.
;    APPLICANT:  Kishore, Ganesh M.
;    APPLICANT:  Padgette, Stephen R.
;    TITLE OF INVENTION:  Glyphosate Tolerant
;    TITLE OF INVENTION:  5-Enolpyruvylshikimate-3-Phosphate Synthases
;    NUMBER OF SEQUENCES:  36
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dennis R. Hoerner, Jr., Monsanto Co. BB4F
;      STREET:  700 Chesterfield Village Parkway
;      CITY:  St. Louis
;      STATE:  Missouri
;      COUNTRY:  USA
;      ZIP:  63198
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/06148A
;      FILING DATE:  19910828
;    , CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/749,611
;      FILING DATE:  28-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/576,537
;      FILING DATE:  31-AUG-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hoerner Jr., Dennis R.
;      REGISTRATION NUMBER:  30,914
;      REFERENCE/DOCKET NUMBER:  38-21(10535)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314)537-6099
;      TELEFAX:  (314)537-6047
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-06148A-7


  Query Match             83.3%;  Score 1906.5;  DB 5;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 4.7e-168;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:|||:|||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | ||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447


Search completed: August 15, 2002, 13:59:16
Job time: 90 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 14:05:43 ; Search time 51.67 Seconds
(without alignments)
1523.373 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2098	91.7	455	16	Q92SV5	Q92sv5 rhizobium m
2	1855.5	81.1	480	2	Q9AGV2	Q9agv2 brucella ab
3	1783	77.9	452	16	Q98CC1	Q98cc1 rhizobium l
4	1135	49.6	443	16	Q9A2H2	Q9a2h2 caulobacter
5	989.5	43.2	746	16	Q9HZ69	Q9hz69 pseudomonas
6	912.5	39.9	454	16	Q9PB21	Q9pb21 xylella fas
7	884	38.6	431	16	Q9KCA6	Q9kca6 bacillus ha
8	820	35.8	430	16	Q99Z83	Q99z83 streptococc
9	815.5	35.6	428	16	Q92A85	Q92a85 listeria in
10	781.5	34.2	428	2	Q9ANY6	Q9any6 enterococcu
11	610	26.7	432	16	Q99U25	Q99u25 staphylococ
12	470.5	20.6	207	2	Q9RHZ8	Q9rhz8 pseudomonas
13	470.5	20.6	439	17	Q9HQC1	Q9hqc1 halobacteri
14	416.5	18.2	440	2	Q9L213	Q91213 streptomyce
15	401.5	17.5	426	16	Q9KRB0	Q9krb0 vibrio chol
16	379	16.6	427	2	Q93ED4	Q93ed4 yersinia ru
17	351.5	15.4	428	16	Q97KM2	Q97km2 clostridium
18	332.5	14.5	438	2	Q9K4A7	Q9k4a7 streptomyce
19	324.5	14.2	447	16	Q9K9D5	Q9k9d5 bacillus ha
20	306	13.4	516	10	Q946V0	Q946v0 dicliptera
21	302	13.2	521	10	Q9FVP6	Q9fvp6 arabidopsis
22	299	13.1	444	10	O24566	O24566 zea mays (m
23	298	13.0	516	10	Q946U9	Q946u9 dicliptera
24	297.5	13.0	511	10	Q93VK6	Q93vk6 oryza sativ
25	293.5	12.8	445	8	Q95AK0	Q95ak0 eleusine in
26	292.5	12.8	445	8	Q95AK1	Q95ak1 eleusine in
27	285	12.5	433	16	Q9JTT3	Q9jtt3 neisseria m
28	280	12.2	414	17	Q980I5	Q980i5 sulfolobus
29	275	12.0	433	16	Q9JYU1	Q9jyu1 neisseria m
30	261	11.4	408	17	Q96Y91	Q96y91 sulfolobus
31	260.5	11.4	427	17	Q9YC47	Q9yc47 aeropyrum p
32	229	10.0	332	2	Q9ZEQ0	Q9zeq0 actinobacil
33	223	9.7	410	17	Q978S3	Q978s3 thermoplasm
34	216.5	9.5	391	10	O80428	O80428 oryza sativ
35	206.5	9.0	448	2	Q9L1U5	Q9l1u5 streptomyce
36	203.5	8.9	446	2	Q9S0N1	Q9s0n1 streptomyce
37	185.5	8.1	419	16	Q99Z78	Q99z78 streptococc
38	184.5	8.1	417	16	Q9K1Q9	Q9k1q9 neisseria m
39	184	8.0	418	16	Q97DD9	Q97dd9 clostridium
40	183.5	8.0	417	16	Q9JWS7	Q9jws7 neisseria m
41	182	8.0	428	16	Q9K6E5	Q9k6e5 bacillus ha
42	181	7.9	419	2	Q9S5W5	Q9s5w5 escherichia
43	179	7.8	347	10	Q9AT37	Q9at37 lolium rigi
44	179	7.8	423	16	Q927U1	Q927u1 listeria in
45	178.5	7.8	417	2	Q9EXE3	Q9exe3 mycobacteri

ALIGNMENTS

RESULT 1
Q92SV5
ID Q92SV5 PRELIMINARY; PRT; 455 AA.
AC Q92SV5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN (EC
DE 2.5.1.19).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
DR EMBL; AL591783; CAC41690.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 91.7%; Score 2098; DB 16; Length 455;
Best Local Similarity 90.5%; Pred. No. 2.3e-119;
Matches 412; Conservative 18; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

```
Db      1 MSHGSNPRPATARKSSDLKGTLRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db     61 KAMQAMGARIRKEGDTWIIDGVGNGALLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db    121 GDASLTKRPMGRVLDPLREMGVQVKSAEGDRLPVTLRGPKTPNPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db    181 LLAGLNTPGITTVVEPVMTRDHTEKMLQGFGANLTVETDAEGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db    241 VPGDPSSTAFPLVAGLIVPGSDITILNVLMNPTRTGLILTLQEMGANIEVMNKRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db    301 VADLRVRHSELKGVTVPEDRAPSMIDEYPVLAVAAAFAEGTTVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db    361 VADGLKLNGVDCDEGEASLVVRGRPGGKGLGKISGGQVKTHLDHRIAMSFLVMGLASEHP 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
Db    421 VTVDDATMIATSFPEFMGLMTGLGAKIEEAENKAA 455
```

RESULT 2
Q9AGV2
ID Q9AGV2 PRELIMINARY; PRT; 480 AA.
AC Q9AGV2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE 5-ENOLPYRUVYL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).
GN AROA.
OS Brucella abortus.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2308;
RA Gan T., Essenberg R.C.;
RT "Characterization of the aroA gene of Brucella abortus and
RT construction of an aroA mutant.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF326475; AAK27445.1; -.
DR InterPro; IPR001986; EPSP_syntase.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 480 AA; 50634 MW; 3D55323944DA8C91 CRC64;

Query Match 81.1%; Score 1855.5; DB 2; Length 480;
Best Local Similarity 81.4%; Pred. No. 1.2e-104;
Matches 364; Conservative 35; Mismatches 45; Indels 3; Gaps 1;

```
Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db     31 MSHSACPKPATARHSQALTGEIRIPGDKSISTRSFMFGGLASGKTRITGLLEGEDVINTG 90

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db     91 RAMQAMGARIRKEGDVWIINGVGNGCLLQPEAPLDFGNAGTGARLTMGLVGTYDMKTSFI 150
```

```
Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db    151 GDASLSKRPMGRVLNPLREMGVQVEAAEGDRMPLTLIGPRTANPIAYRVPMASAQVKSAV 210

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db    211 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRIVGQGKLTGQTID 270

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db    271 VPGDPSSTAFPLVPALLVEGSEVTIRNVLMNPTRTGLILTLQEMGADIEIIDPRLAGGED 330

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db    331 VADLRVKASKLKGVVVPPERAPSMIDEYPVLAIXASFAEGETVMDGLDELRVKESDRLAA 390

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db    391 VARGLEANGVDCTEGEMSLTVRGRPGGKGLG---GGTVATHLDHRIAMSFLVMGLASEKP 447

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKI 447
Db    448 VTVDDSTMIATSFPEFMGMMAGLGAKI 474
```

RESULT 3
Q98CC1
ID Q98CC1 PRELIMINARY; PRT; 452 AA.
AC Q98CC1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN MLL5213.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51700.1; -.
DR InterPro; IPR001986; EPSP_syntase.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 452 AA; 47455 MW; 2B52983E3523B938 CRC64;

Query Match 77.9%; Score 1783; DB 16; Length 452;
Best Local Similarity 77.8%; Pred. No. 2.7e-100;
Matches 351; Conservative 33; Mismatches 63; Indels 4; Gaps 2;

```
Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db      1 MSHAAAAKPATARKSQALSGTARVPGDKSISHRSFMFGGLASGETRITGLLEGEDVMRTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db     61 AAMKAMGAHIEKRGAEWVIRGTGNGALLQPEGPLDFGNAGTGSRLTMGLVGTYDMETTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQV-KSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSA 179
```

```
Db    121 GDASLSGRPMGRVLEPLRQMGVQVLKATPGDRMPITLHGPKHAAPITYRVPMASAQVKSA 180

Qy    180 VLLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVI 239
          ||||||||||||||||||:||||||||||:||||||:||||  ||| | :||:|:|||| |
Db    181 VLLAGLNTPGITTVIEPVMTRDHTEKMLKGFGANLSVETDERGVRHIFIEGQGRLTGQTI 240

Qy    240 DVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGE 299
          |||||||  |||||||:|||||:|| ||||||||||||:|||||||  |:::||| ||||
Db    241 AVPGDPSSAGFPLVAALIVPGSDITIENVLMNPTRTGLLLTLQEMGGQIDILNPRNAGGE 300

Qy    300 DVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLS 359
          |||||||| | |||| || :|||||||||:|||||:|||| |:| ||||||||||||||
Db    301 DVADLRVRYSELKGVAVPPERAPSMIDEYPVLAVAASFAEGETLMQGLEELRVKESDRLS 360

Qy    360 AVANGLKLNGVDCDEGETSLVVRGRPDGKGLG---NASGAAVATHLDHRIAMSFLVMGLV 416
          |||||||||||||| ||| || ||||| |||||    |      | ||||||||||||||||
Db    361 AVANGLKLNGVDCTEGEASLAVRGRPGGKGLGGHPNGLDTTVQTHLDHRIAMSFLVMGLA 420

Qy    417 SENPVTVDDATMIATSFPEFMDLMAGLGAKI 447
          :| |||:||  |||||||||| || ||||:|
Db    421 TEKPVTIDDQAMIATSFPEFMGLMTGLGAEI 451
```

RESULT 4
Q9A2H2
ID Q9A2H2 PRELIMINARY; PRT; 443 AA.
AC Q9A2H2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN CC3589.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE006017; AAK25551.1; -.
DR TIGR; CC3589; -.
DR InterPro; IPR001986; EPSP_syntase.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
KW Transferase; Complete proteome.
SQ SEQUENCE 443 AA; 46075 MW; 45E8A1463E10B6EC CRC64;

Query Match 49.6%; Score 1135; DB 16; Length 443;
Best Local Similarity 55.8%; Pred. No. 4.5e-61;
Matches 251; Conservative 45; Mismatches 138; Indels 16; Gaps 6;

```
Qy      2 LHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGK 61
          | |  | |  |     | | || |||||||||| : | ||:| | : |||||:||: | :
Db      3 LAGLKSAPGGA-----LRGIVRAPGDKSISHRSMILGALATGTTTVEGLLEGDDVLATAR 57

Qy     62 AMQAMGARIRKEG-DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          |||| |||| :||   | |:  | ||   |   :| ||| || || ||    :   :||
Db     58 AMQAFGARIEREGVGRWRIE--GKGGFEEPVDVIDCGNAGTGVRLIMGAAAGFAMCATFT 115

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          || ||  |||||||:||  ||      |  |||:||:|      : | :||||||||||
```

```
Db    116 GDQSLRGRPMGRVLDPLARMGATWLGRDKGRLPLTLKGGNL-RGLNYTLPMASAQVKSAV 174

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRT---IRLEGRGKLTGQ 237
          |||||:  |   ||||  ||||||:||: ||| : ||    | :|   :||    ||||
Db    175 LLAGLHAEGGVEVIEPEATRDHTERMLRAFGAEVIVEDRKAGDKTFRHVRLPEGQKLTGT 234

Qy    238 VIDVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAG 297
          : ||||||| ||||||||:||||:||:  |::|  ||||  ||||||||: : | |:|
Db    235 HVAVPGDPSSAAFPLVAALIVPGSEVTVEGVMLNELRTGLFTTLQEMGADLVISNVRVAS 294

Qy    298 GEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDR 357
          ||:| |:  | | |||| || :|||||||||||||||||| | ||| |: |:|||||||
Db    295 GEEVGDITARYSQLKGVVVPPERAPSMIDEYPILAVAAAFASGETVMRGVGEMRVKESDR 354

Qy    358 LSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVS 417
          :|  |||||  ||   |    :|     | |      || | || ||||||| |:::|: :
Db    355 ISLTANGLKACGVQVVEEPEGFIV----TGTGQPPKGGATVVTHGDHRIAMSHLILGMAA 410

Qy    418 ENPVTVDDATMIATSFPEFMDLMAGLGAKI 447
          :  | ||:  |||||| | ||| |||| :
Db    411 QAEVAVDEPGMIATSFPGFADLMRGLGATL 440
```

RESULT 5
Q9HZ69
ID Q9HZ69 PRELIMINARY; PRT; 746 AA.
AC Q9HZ69;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE STILL FRAMESHIFT 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE
DE PREPHENATE DEHYDROGENASE.
GN PA3164.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004740; AAG06552.1; -.
DR InterPro; IPR001986; EPSP_syntase.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR003099; PDH.
DR Pfam; PF00275; EPSP_syntase; 1.
DR Pfam; PF02153; PDH; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 746 AA; 79320 MW; C2974B4BBF539E3A CRC64;

Query Match 43.2%; Score 989.5; DB 16; Length 746;
Best Local Similarity 48.5%; Pred. No. 6e-52;
Matches 214; Conservative 62; Mismatches 152; Indels 13; Gaps 3;

```
Qy     12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
          |:    ||||:|:|||||||||| | | || | | : | ||||| : | :| : ||  |
Db    317 AQPGGSLSGTIRVPGDKSISHRSIMLGSLAEGTTEVEGFLEGEDALATIQAFRDMGVVIE 376

Qy     72 -KEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
           :    : |||  || ||  |:  ||: |  ||  ||:    ||||  |||||:||||
```

```
Db    377 GPQNGRVTVHGVGLHGLKAPPGPIYLGNSGTSMRLLSGLLAAQPFDSTLTGDASLSKRPM 436

Qy    131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
          ||  |||||| :::    | |:|:|| :  | : | :|||||||| :|||||   |
Db    437 NRVAKPLREMGAVIETGPEGRPPMTIRGGQRLTGMHYDMPMASAQVKSCLLLAGLYAAGE 496

Qy    191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
          |:| ||  ||||||:||:|||  : ||       | ::|   ||:   |:|| | || ||
Db    497 TSVTEPAPTRDHTERMLRGFGYPVVVEGS-----TAKVESGHKLSATHIEVPADISSAAF 551

Qy    251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
          |||| :  ||:: : :| :|||| |:|  |: || |: : | |  ||| |||:|||||:
Db    552 FLVAASIAEGSELVLQHVGINPTRVGVIEILRLMGGDLSLENQREVGGEPVADIRVRSAR 611

Qy    311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
          |||: :|||  |  |||:|:| |||| ||| ||: | ||||||||||:  :|:|||  ||
Db    612 LKGIDIPEDLVPLAIDEFPVLFVAAACAEGRTVLRGAEELRVKESDRIQVMADGLKALGV 671

Qy    371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
          :    :|: |       |   |  | |||||||| |  | :  |: : |   :|
Db    672 KAEPTPDGIVIEG-------GAFGGGEVWAHGDHRIAMSFSVASLRASGPIRIHDCANVA 724

Qy    431 TSFPEFMDLMAGLGAKIELSD 451
          |||| |: | |  | :: : :
Db    725 TSFPNFLALCAQTGIRVAVEN 745
```

```
RESULT   6
Q9PB21
ID   Q9PB21      PRELIMINARY;      PRT;   454 AA.
AC   Q9PB21;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   XF2324.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zaha M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AF004043; AAF85123.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   454 AA;  48266 MW;  45CCF074E6C0BA57 CRC64;
```

```
  Query Match             39.9%;  Score 912.5;  DB 16;  Length 454;
  Best Local Similarity   46.9%;  Pred. No. 1.4e-47;
  Matches  206;  Conservative   60;  Mismatches  158;  Indels   15;  Gaps    4;

Qy     12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
          | : : | | : ||||||||||: ||  || | :|| | || ||  :| : :  :| ||
Db     23 AHQGTPLHGVLSIPGDKSISHRAVMFAALADGTSRIDGFLEAEDTCSTAEILARLGVRIE 82

Qy     72 KEGDTW-IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
             |  |: |||  || |   ||| ||| || ||  ||:    |||  :|||||:||||
Db     83 TPLSTQRIVHGVGVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVGDASLSKRPM 142

Qy    131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
          || :|| :|| :: : |    |: : | :    | :  |:||||:|||||||||
Db    143 RRVTDPLSQMGARIDTSDDGTPPLRIYGGQLLHGIDFISPVASAQIKSAVLLAGLYARNE 202

Qy    191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
          | | ||  |||:||:||  || :: | |        || |  :|    | :| | || ||
Db    203 TVVREPHPTRDYTERMLTAFGVDIDVSTGC-----ARLRGGQRLCATDITIPADFSSAAF 257

Qy    251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
          |||| ::|||:|:  | :|| | ||:  |: |||:|    |    ||| | ||||| :
Db    258 YLVAASVIPGSDITLRAVGLNPRRIGLLTVLRLMGANIVESNRHEQGGEPVVDLRVRYAP 317

Qy    311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
          |:|  ||||    ||||:| | |||| ||| ||::|  ||||||||||:|:  ||:: ||
Db    318 LQGTRVPEDLVADMIDEFPALFVAAAAAEGQTVVSGAAELRVKESDRLAAMVTGLRVLGV 377

Qy    371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
            ||      : | | | |  |: |       ||||||:| : | :| : | ::|   :|
Db    378 QVDETADGATIHGGPIGHGTINSHG-------DHRIAMAFSIAGQLSVSTVRIEDVANVA 430

Qy    431 TSFPEFMDL--MAGLGAKI 447
          ||||::  |   || | ::
Db    431 TSFPDYETLARSAGFGLEV 449
```

```
RESULT   7
Q9KCA6
ID   Q9KCA6      PRELIMINARY;      PRT;   431 AA.
AC   Q9KCA6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE   ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)
DE   (EPSPS).
GN   AROE OR BH1667.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =
CC       ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
```

```
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC       AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR   EMBL; AP001512; BAB05386.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ   SEQUENCE   431 AA;  45485 MW;  12F4FFBE7BA0743D CRC64;

  Query Match           38.6%;  Score 884;  DB 16;  Length 431;
  Best Local Similarity   44.2%;  Pred. No. 6.9e-46;
  Matches  188;  Conservative   72;  Mismatches  153;  Indels   12;  Gaps    5;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           : || ||:::|||||||||: ||| || | | : | | | | ::|    | :|  | :
Db      10 AKGLKGTIKVPGDKSISHRAVMFGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |  ||  |   || ||: |  || :|::    | |  ||| |: |||| ||
Db      70 ERVTVKGKGWDGLREPSDILDVGNSGTTTRLILGILSTLPFHSVIIGDESIGKRPMKRVT 129

Qy     135 NPLREMGVQVKSED-GDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || |:   | |:  |:::|| :    | :  |:||||:|||:|||   | |:|
Db     130 EPLKSMGAQIDGRDHGNLTPLSIRGGQL-KGIDFHSPVASAQMKSAILLAGLRAEGKTSV 188

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           ||  ||||||:||: || |:    : ||: |: :||   |||| : |||| || || ||
Db     189 TEPAKTRDHTERMLEAFGVNI----EKDGL-TVSIEGGQMLTGQHVVVPGDISSAAFFLV 243

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
           |  :|| | :|: || :|||| |::  |::||| : : | |: ||| |||| : :| |:|
Db     244 AGAMVPHSRITLTNVGINPTRAGILEVLKQMGATLAMENERVQGGEPVADLTIETSVLQG 303

Qy     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373
           | :  |  | :||| ||:|| |  | | ||:   |||:|||::|:  | : |   |
Db     304 VEIGGDIIPRLIDEIPIIAVLATQASGRTVIKDAEELKVKETNRIDTVVSELTKLGASIH 363

Qy     374 EGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSF 433
           :  ::: |    ||     |  |::| |||| |:  :  |::| ||||:    || |:
Db     364 ATDDGMIIEGPTPLKG-----GVTVSSHGDHRIGMAMAIAALLAEKPVTVEGTEAIAVSY 418

Qy     434 PEFMD 438
           | | |
Db     419 PSFFD 423


RESULT   8
Q99Z83
ID   Q99Z83      PRELIMINARY;      PRT;   430 AA.
AC   Q99Z83;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE
DE   (EC 2.5.1.19).
GN   AROA OR SPY1352.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006573; AAK34180.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   430 AA;  46692 MW;  3273C6B39020FB61 CRC64;

  Query Match           35.8%;  Score 820;  DB 16;  Length 430;
  Best Local Similarity   43.8%;  Pred. No. 5.1e-42;
  Matches  189;  Conservative   76;  Mismatches  141;  Indels   26;  Gaps    9;

Qy      13 RKSSG-LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
           | ::| | ||:::|||||||||: : | :| |||| |||:||||::| :| : :| ||
Db       7 RTNAGPLQGTIQVPGDKSISHRAVILGAVAKGETRVKGLLKGEDVLSTIQAFRNLGVRIE 66

Qy      72 KEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMG 131
           :: |  :|:| |  || ||   |: ||: |  ||  ||:    |    ||| ||:||||
Db      67 EKDDQLVIEGQGFQGLNAPCQTLNMGNSGTSMRLIAGLLAGQPFSVKMIGDESLSKRPMD 126

Qy     132 RVLNPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
           |:: ||::|||::  | |    |: |:| :   |||| :|::|||||||:||| |   |
Db     127 RIVYPLKQMGVEISGETDRQFPPLQLQGNRNLQPITYTLPISSAQVKSAILLAALQAKGT 186

Qy     191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
           | |:|  :||:|||:|:| ||  | |    || | | | |  :|| | | |||| || ||
Db     187 TQVVEKEITRNHTEEMIQQFGGRLIV----DGKR-ITLVGPQQLTAQEITVPGDISSAAF 241

Qy     251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADI--EVINPRLAGGEDVADLRVRS 308
           |||  |::|||:: : || :||||||::  :::||| |  | :| :    | |  :||
Db     242 WLVAGLIIPGSELLLKNVGVNPTRTGILEVVEKMGAQIVYEDMNKK----EQVTSIRVVY 297

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
           | :||  :     | :||| ||:|: |  |:| | :   :||||||:||:  | : |
Db     298 SNMKGTIISGGLIPRLIDELPIIALLATQAQGTTCIKDAQELRVKETDRIQVVTDILNSM 357

Qy     369 GVDCDEGETSLVVRGRPDG---KGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVD 424
           | :          ::   ||   ||     ||  :|: |||| |   :   || :  | :|
Db     358 GAN---------IKATADGMIIKGPTVLYGANTSTYGDHRIGMMTAIAALLVKQGQVHLD 408

Qy     425 DATMIATSFPEF 436
               | ||:| |
Db     409 KEEAIMTSYPTF 420


RESULT   9
Q92A85
ID   Q92A85      PRELIMINARY;      PRT;   428 AA.
AC   Q92A85;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AROE PROTEIN.
GN   AROE OR LIN2037.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
```

```
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596170; CAC97267.1; -.
DR   ListiList; LIN02037; -.
KW   Complete proteome.
SQ   SEQUENCE   428 AA;  45994 MW;  157B48C091A68FEB CRC64;
```

```
  Query Match             35.6%;  Score 815.5;  DB 16;  Length 428;
  Best Local Similarity   40.8%;  Pred. No. 9.5e-42;
  Matches  173;  Conservative   83;  Mismatches  155;  Indels   13;  Gaps    5;

Qy      17 GLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT 76
           || | : :|||||:|||| ||| :| |:| |   |  :| : | || :|:| :| :  :
Db       9 GLVGEITVPGDKSMSHRSIMFGAIAEGKTVIRHFLRADDCLGTIKAFKALGVKIEETDEE 68

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNP 136
           |: | |: ||   | ||| ||: |  || ||::   |||:  :|| |: |||| ||: |
Db      69 IIVHGTGSDGLKQAEGPLDIGNSGTTIRLMMGILAGRDFDTVILGDESIAKRPMNRVMLP 128

Qy     137 LREMGVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
           |:||| ::  :||    |::: | ::   : | :|:|||||||:: | |   | | : |
Db     129 LQEMGAKMHGKDGSEFAPISIIGNQSLKRMEYHMPVASAQVKSAIIFAALQAEGETIIHE 188

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
           |||||| |:: ||     | : ||: |||::|  |  || : |||| || || :||
Db     189 KEKTRDHTEHMIRQFGG----EIEMDGL-TIRVKGGQKFIGQEMTVPGDVSSAAFFIVAG 243

Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
           |: |||:: : :| :||||||:   :::||  : | :   : |:    : |:|| |||
Db     244 LITPGSEIELTHVGLNPTRTGIFDVVEQMGGSLVVKDSSRSTGKLAGTVVVKSSELKGTE 303

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
           :  |  | :||| |::|: |  ||| |::    ||:|||::|: |||| |   | |
Db     304 IGGDIIPRLIDEIPVIALLATQAEGTTIIKDAAELKVKETNRIDAVANELNKMGADITPT 363

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSFP 434
           |  |::||:           | | :: |||| |   :   || :  | :| |  :: |:|
Db     364 EDGLIIRGKTP------LHAANVTSYGDHRIGMMLQIAALLVEDGDVELDRAEAVSVSYP 417

Qy     435 EFMD 438
           | :
Db     418 TFFE 421
```

```
ESULT  10
ANY6
D   Q9ANY6      PRELIMINARY;      PRT;   428 AA.
AC   Q9ANY6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19).
GN   AROA.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC   Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OG1RF;
RA   Huycke M.M., Shepard L., Joyce W., Wise P., Moore D.R., Gilmore M.S.;
RT   "Forme fruste respiration by Enterococcus faecalis produces
RT   extracellular superoxide.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF318277; AAG53678.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE;, PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase.
SQ   SEQUENCE   428 AA;  45715 MW;  9E0F4FE4A893CA95 CRC64;
```

```
  Query Match             34.2%;  Score 781.5;  DB 2;  Length 428;
  Best Local Similarity   40.5%;  Pred. No. 1.1e-39;
  Matches  173;  Conservative   79;  Mismatches  152;  Indels   23;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
           | ||: :| |||||||| ||| ::||:| ||  | ||| ::|  | :::|  |  :| |
Db      12 LQGTLMVPSDKSISHRSIMFGAISSGKTTITNFLRGEDCLSTLAAFRSLGVNIEDDGTTI 71

Qy      78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
           ::| |  ||   :  :| ||: |  || :|::    |::   ||||: |||| ||: ||
Db      72 TVEGRGFAGLKKAKNTIDVGNSGTTIRLMLGILAGCPFETRLAGDASIAKRPMNRVMLPL 131

Qy     138 REMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196
           :|| : :  :  :  |:::|| :   || | :|:|||||||:| | |   | : |:|
Db     132 NQMGAECQGVQQTEFPPISIRGTQNLQPIDYTMPVASAQVKSAILFAALQAEGTSVVVEK 191

Qy     197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
           ||||||:|:: ||  | |    || : | | |  :|||| : |||| || || ||| |
Db     192 EKTRDHTEEMIRQFGGTLEV----DG-KKIMLTGPQQLTGQNVVVPGDISSAAFFLVAGL 246

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVA----DLRVRSSTLK 312
           :|| |:: : || :| ||||::  :: ||  : ::|      || |    || |:|
Db     247 VVPDSEILLKNVGLNQTRTGILDVIKNMGGSVTILN------EDEANHSGDLLVKTSQLT 300

Qy     313 GVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDC 372
           :      | :||| ||:|: |  | | |::   |||:|||::|: |||  | : | |
Db     301 ATEIGGAIIPRLIDELPIIALLATQATGTTIIRDAEELKVKETNRIDAVAKELTILGADI 360

Qy     373 DEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIAT 431
           :  |:: |        :  |  | :: |||| |   :   || |  | :| |  ::
Db     361 TPTDDGLIIHGPT------SLHGGRVTSYGDHRIGMMLQIAALLVKEGTVELDKAEAVSV 414

Qy     432 SFPEFMD 438
           |:| | |
Db     415 SYPAFFD 421
```

```
RESULT  11
Q99U25
ID   Q99U25      PRELIMINARY;      PRT;   432 AA.
AC   Q99U25;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   AROA OR SA1297 OR SAV1464.
OS   Staphylococcus aureus (strain N315), and
OS   Staphylococcus aureus (strain Mu50).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=158879, 158878;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003134; BAB42557.1; -.
DR   EMBL; AP003362; BAB57626.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
```

```
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   432 AA;  47068 MW;  A42102057AD15C72 CRC64;


  Query Match             26.7%;  Score 610;  DB 16;  Length 432;
  Best Local Similarity   34.0%;  Pred. No. 2.7e-29;
  Matches  146;  Conservative   80;  Mismatches  181;  Indels   22;  Gaps    6;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           |  | | : :|||||::||: |   || | : |   | |||   |    : :|  |::: 
Db      10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYKPLLGEDCRRTMDIFRLLGVDIKEDE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           |  :::  |       |   | ||: |  ||  ||:     :|   || |: |||| |||
Db      70 DKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLAGLLSGLGIESVLSGDVSIGKRPMDRVL 129

Qy     135 NPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
           ||: |   ::  : :  |: :: |     | |:: :|||||||:| | | :   | :
Db     130 RPLKLMDANIEGIEDNYTPLIIK-PSVIKGINYQMEVASAQVKSAILFASLFSNDTTVIK 188

Qy     195 EPIMTRDHTEKMLQGF-----GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249
           |  ::|:||| | : |        |:: |  | :: |:              |||| || |
Db     189 ELDVSRNHTETMFRHFNIPIEAERLSITTTPDAIQHIKPAD--------FHVPGDISSAA 240

Qy     250 FPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVR-S 308
           | :|||: | ||||| || :||||:|:|  :::|| :|:: | :  | |  | :|:: :
Db     241 FFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMGGNIQLFN-QTTGAEPTASIRIQYT 299

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
           |: :|:  :  |  ||| |::|:    | | : :   |||:|||::|:   |: | |
Db     300 PMLQPITIEGELVPKAIDELPVIALLCTQAVGTSTIKDAEELKVKETNRIDTTADMLNLL 359

Qy     369 GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATM 428
           | :       |::            : | | :  |||| |   |  |:|  || :
Db     360 GFELQPTNDGLIIHPSE------FKTNATVDSLTDHRIGMMLAVASLLSSEPVKIKQFDA 413

Qy     429 IATSFPEFM 437
           :  ||| |:
Db     414 VNVSFPGFL 422


RESULT  12
Q9RHZ8
ID   Q9RHZ8      PRELIMINARY;      PRT;   207 AA.
AC   Q9RHZ8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   5-ENOLPYRUVYLSHIKMATE 3-P SYNTHASE (FRAGMENT).
GN   AROF.
OS   Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=316;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JM300;
RX   MEDLINE=92013931; PubMed=1919506;
RA   Fischer R.S., Zhao G., Jensen R.A.;
RT   "Cloning, sequencing, and expression of the P-protein gene (pheA) of
RT   Pseudomonas stutzeri in Escherichia coli: implications for
RT   evolutionary relationships in phenylalanine biosynthesis.";
RL   J. Gen. Microbiol. 137:1293-1301(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JM300;
RX   MEDLINE=99298294; PubMed=10368439;
RA   Xie G., Bonner C.A., Jensen R.A.;
RT   "A probable mixed-function supraoperon in Pseudomonas exhibits gene
RT   organization features of both intergenomic conservation and gene
RT   shuffling.";
RL   J. Mol. Evol. 49:108-121(1999).
DR   EMBL; AF038578; AAD47363.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
FT   NON_TER     207     207
SQ   SEQUENCE   207 AA;  21780 MW;  7231191C72A21D6B CRC64;


  Query Match             20.6%;  Score 470.5;  DB 2;  Length 207;
  Best Local Similarity   51.5%;  Pred. No. 2.7e-21;
  Matches  103;  Conservative   22;  Mismatches   70;  Indels    5;  Gaps    2;

Qy      12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
           |      |:| :|:|||||||||||| | | || | |   | ||||| : | :| : ||  |
Db      10 ANPGGSLTGQLRVPGDKSISHRSIMLGSLAEGTTEXEGFLEGEDALATXQAFRXMGVVI- 68

Qy      72 KEG---DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKR 128
           ||        : |||  || ||  |:  ||: |  ||  ||:    ||:|  ||||||||
Db      69 -EGPHQGRVTVHGVGLHGLQAPPGPIYLGNSGTSMRLLAGLLAAQPFDTTLSGDASLTKR 127

Qy     129 PMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTP 188
           || ||  ||||||  :::    | |:|:|| |  : : | :|||||||| :|||||
Db     128 PMNRVAKPLREMGAVIETAAEGRPPLTIRGGKKLSGMHYDMPMASAQVKSCLLLAGLYAA 187

Qy     189 GITTVIEPIMTRDHTEKMLQ 208
           | |:| ||  ||||||:| |
Db     188 GKTSVTEPAPTRDHTERMQQ 207


RESULT  13
Q9HQC1
ID   Q9HQC1      PRELIMINARY;      PRT;   439 AA.
AC   Q9HQC1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   PSC OR VNG1232G.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005049; AAG19594.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   439 AA;  44333 MW;  48CAD75A1F0ECB89 CRC64;


  Query Match             20.6%;  Score 470.5;  DB 17;  Length 439;
  Best Local Similarity   31.3%;  Pred. No. 7.6e-21;
  Matches  139;  Conservative   76;  Mismatches  190;  Indels   39;  Gaps   13;
```

```
Qy      16 SGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75
Db      17 SRVRGRARAPPSKSYTHRALLAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG 76

Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db      77 DWVVTGFGSRPAI-PDAVIDCANSGTTMRLVTAAAALADGTTVLTGDESLRARPHGPLLD 135

Qy     136 PLREMGVQVKSEDGD-RLPVTLRGPKTPTPITYRVP-MASAQVKSAVLLAGLNT-----P 188
Db     136 ALSGLGGTARSTRGNGQAPLVVDGPVSGGSVA--LPGDVSSQFVTALLMAGAVTETGIET 193

Qy     189 GITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRG----KLTGQVIDVPGD 244
Db     194 DLTTELKSAPYVDITLDVLDAFGVGAS-ET-AAGYRV-----RGGQAYAPSGAEYAVPGD 246

Qy     245 PSSTAFPLVAALLVPGSDVTILNVLMNPTRTG---LILTLQEMGADIEVINPRLAGGEDV 301
Db     247 FSSASYLLAAGALAAADGAAVVVEGMHPSAQGDAAIVDVLERMGADID-------WDTES 299

Qy     302 ADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAV 361
Db     300 GVITVQRSELSGVEVGVADTPDLL---PTIAVLGAAADGTTRITDAEHVRYKETDRVAAM 356

Qy     362 ANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPV 421
Db     357 AESLSKLGASVEERPDELVVRG-----GDTELSGASVDGRGDHRLVMALAVAGLVADGET 411

Qy     422 TVDDATMIATSFPEFMDLMAGLGA 445
Db     412 TIAGSEHVDVSFPDFFEVLAGLGA 435
```

RESULT 14
Q9L213
ID Q9L213 PRELIMINARY; PRT; 440 AA.
AC Q9L213;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN AROA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL138598; CAB71266.1; -.
DR InterPro; IPR001986; EPSP_syntase.
DR InterPro; IPR000408; RCC1.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW Transferase.

SQ SEQUENCE 440 AA; 46529 MW; 77B720F81398EB0D CRC64;

Query Match 18.2%; Score 416.5; DB 2; Length 440;
Best Local Similarity 28.0%; Pred. No. 1.4e-17;
Matches 128; Conservative 78; Mismatches 198; Indels 53; Gaps 13;

```
Qy      11 TARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMG--- 67
Db      15 TIRTLTGFDETVRVLGSKSYTNRYLAIASLSGQETVIDNALLSDDTVYFSRAIETFGHVT 74

Qy      68 -------ARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      75 CDIDHATARIR-------VTPTGR-PMRAPSEDIFVGGAGTPLRFLISMAGHADGTTIIT 126

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGD-RLPVTLRGPKTPTPITYRVPMASAQVKSA 179
Db     127 GNARMQERPMGDLLKALPALGVDATAVRGNGSPPVRVVGGSFKGGATSISGAVSSQFTSS 186

Qy     180 VLLAGLNTPGIT--TVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQ 237
Db     187 LIINALRAQTDTEITISDDLVSKPYVEMTLAGL-AEMGVSVDRDGYRRFTVPSGQQARGG 245

Qy     238 VIDVPGDPSSTAFPLVAA------LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVI 291
Db     246 QVTVEPDASGMSYFLAAAAILQSRVVIPG-----IGAGSHQGDVHLVQALERMGCRTEV- 299

Qy     292 NPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELR 351
Db     300 ------GDD--SITVTGGPLRGIDIDMEAMPDVV---PSLAIVAAYAEGTTRITNIASLR 348

Qy     352 VKESDRLSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFL 411
Db     349 VKECDRIAAVTTELRKMGIDVEEHSDAMYITG-------GTPHGAVIDTYDDHRIAMTFA 401

Qy     412 VMGLVSENPVTVDDATMIATSFPEFMDLMAGLGAKIE 448
Db     402 IGGLRTEG-VVIKDPGCVAKSFPAFWQTLDTLHPDLE 437
```

RESULT 15
Q9KRB0
ID Q9KRB0 PRELIMINARY; PRT; 426 AA.
AC Q9KRB0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)
DE (EPSPS).
GN VC1732.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =
CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

```
DR   EMBL; AE004251; AAF94882.1; -.
DR   TIGR; VC1732; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ   SEQUENCE   426 AA;  46101 MW;  38852D6483BFE1C3 CRC64;


  Query Match             17.5%;  Score 401.5;  DB 16;  Length 426;
  Best Local Similarity   28.0%;  Pred. No. 1.1e-16;
  Matches  125;  Conservative   87;  Mismatches  180;  Indels   55;  Gaps   15;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD-- 75
           :|| | :|| ||:|:|: :   |||| ||:| ||: :|: :   |:  :|   |   |
Db      12 ISGEVNLPGSKSVSNRALLLAALASGTTRLTNLLDSDDIRHMLNALTKLGVNYRLSADKT 71

Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
           |  ::|:|          |  ||| |  |     : :   |    |:  : :||:| :::
Db      72 TCEVEGLGQAFHTTQPLELFLGNAGTAMRPLAAALCLGQGDYVLTGEPRMKERPIGHLVD 131

Qy     136 PLREMGVQVKSEDGDRLP-------------VTLRGPKTP---TPITYRVPMASAQVKSA 179
            ||: | |::  : :  |              ||: |  :    |      |:|  :|
Db     132 ALRQAGAQIEYLEQENFPPLRIQGTGLQAGTVTIDGSISSQFLTAFLMSAPLAQGKVTIK 191

Qy     180 VLLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLT-GQV 238
           ::   :: | |          | |  ::: ||  :    : |    :   |:  :: ||
Db     192 IVGELVSKPYI----------DITLHIMEQFGVQV---INHDYQEFVIPAGQSYVSPGQF 238

Qy     239 IDVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLIL--TLQEMGADIEVINPRLA 296
           : | || || ::  :||  : | :| :  :  |  :  :     |::||| ||
Db     239 L-VEGDASSASY-FLAAAAIKGGEVKVTGIGKNSIQGDIQFADALEKMGAQIE------- 289

Qy     297 GGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESD 356
           |:|   :  |   | | :  :  |   |    :|  | ||:| | :  :   ||||:|
Db     290 WGDDY--VIARRGELNAVDLDFNHIP---DAAMTIATTALFAKGTTAIRNVYNWRVKETD 344

Qy     357 RLSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLV 416
           ||:|:|  |:  |   :|||  :|:   |  | :     ||: |: |||:|| | :: |
Db     345 RLAAMATELRKVGATVEEGEDFIVI--TPPTKLI----HAAIDTYDDHRMAMCFSLVAL- 397

Qy     417 SENPVTVDDATMIATSFPEFMDLMAGL 443
           |: |||::|    : :||:: |  | |
Db     398 SDTPVTINDPKCTSKTFPDYFDKFAQL 424


rch completed: August 15, 2002, 14:05:44
Job time: 413 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:47 ; Search time 56.25 Seconds
(without alignments)
898.464 Million cell updates/sec

Title: US-09-464-099A-70
Perfect score: 2288
Sequence: 1 MLHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2288	100.0	455	13	AAR26449	CP4-EPSPS. Synthe
2	2288	100.0	455	19	AAW39426	CP4-EPSPS protein.
3	2288	100.0	527	22	AAM52214	EPSPS SEQ ID NO 3.
4	2282	99.7	455	13	AAR22300	Class II EPSPS enz
5	2282	99.7	455	18	AAW34683	Class II EPSP synt
6	2282	99.7	455	18	AAW24474	Class II EPSPS for
7	2282	99.7	455	19	AAW71609	Agrobacterium sp.
8	2282	99.7	455	22	AAE05053	Agrobacterium sp.
9	1900.5	83.1	449	13	AAR22301	Class II EPSPS enz
10	1900.5	83.1	449	13	AAR22302	Class II EPSPS enz
11	1900.5	83.1	449	18	AAW34684	Class II EPSP synt
12	1900.5	83.1	449	18	AAW34685	Class II EPSP synt
13	1900.5	83.1	449	18	AAW24479	Class II EPSPS use
14	1900.5	83.1	449	18	AAW24480	Class II EPSPS use
15	1900.5	83.1	449	19	AAW71611	Pseudomonas sp. st
16	1900.5	83.1	449	19	AAW71610	Achromobacter sp.
17	1900.5	83.1	449	22	AAE05054	Achromobacter sp.
18	1900.5	83.1	449	22	AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18	AAW34688	Class II EPSP synt
20	867.5	37.9	447	18	AAW24487	Class II EPSPS for
21	867.5	37.9	447	19	AAW71619	Synechocystis sp.
22	867.5	37.9	447	22	AAE05072	Synechocystis sp.
23	806	35.2	443	18	AAW34689	Class II EPSP synt
24	806	35.2	443	18	AAW24488	Class II EPSPS for
25	806	35.2	443	19	AAW71620	Dichelobacter nodo
26	806	35.2	443	22	AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22	AAB48177	S. pneumoniae aroA
28	803.5	35.1	427	22	AAB37094	Streptococcus pneu
29	803.5	35.1	427	22	AAB48500	Streptococcus pneu
30	798.5	34.9	427	20	AAW97388	5-enolpyruvylshiki
31	798.5	34.9	427	21	AAY56504	Streptococcus pneu
32	798.5	34.9	427	22	AAB31158	A 5-enolpyruvoylsh
33	785	34.3	428	18	AAW34686	Class II EPSP synt
34	785	34.3	428	18	AAW24481	Class II EPSPS use
35	785	34.3	428	19	AAW71617	Bacillus subtilis
36	785	34.3	428	22	AAE05070	Bacillus subtilis
37	704.5	30.8	415	20	AAW97389	Protein encoded by
38	704.5	30.8	415	21	AAY56505	Streptococcus pneu
39	704.5	30.8	415	22	AAB31159	A 5-enolpyruvoylsh
40	704.5	30.8	415	22	AAB48178	S. pneumoniae aroA
41	704.5	30.8	415	22	AAB37095	Streptococcus pneu
42	704.5	30.8	415	22	AAB48501	Protein encoded by
43	577	25.2	430	18	AAW34687	Class II EPSP synt
44	577	25.2	430	18	AAW24482	Class II EPSPS use
45	577	25.2	430	19	AAW71618	Staphylococcus aur

ALIGNMENTS

RESULT 1
AAR26449
ID AAR26449 standard; Protein; 455 AA.
XX
AC AAR26449;
XX
DT 28-JAN-1993 (first entry)
XX
DE CP4-EPSPS.
XX
KW pMON11030; CTP2; CP4; EPSPS; chloroplast transit peptide;
KW 5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;
KW Pseudomonas chloroaphis 6G5.
XX
OS Synthetic.
XX
PN WO9212249-A.
XX
PD 23-JUL-1992.
XX
PF 17-DEC-1991; 91WO-US09437.
XX
PR 26-DEC-1990; 90US-0632440.
XX
PA (MONS) MONSANTO CO.
XX
PI Kishore GM, Klee HJ;
XX
DR WPI; 1992-284334/34.
DR N-PSDB; AAQ27201.
XX
PT Delaying fruit ripening and senescence in plants - by controlling
PT ethylene prodn., pref. by expression of
PT 1-amino:cyclopropane-1-carboxylic acid deaminase

```
XX
PS   Disclosure; Page 64-66; 110pp; English.
XX
CC   The sequences given in AAR26448 and AAR26449 are encoded by genes which
CC   were used in the construction of an expression plasmid pMON11030 which
CC   was used to transform petunia plants.  This plasmid contained the genes
CC   for chloroplast transit peptide (CTP2) (AAR26448) and the CP4 synthetic
CC   5-enolpyruvyl-3-shikimate phosphate synthetase (EPSPS) gene (AAR26449)
CC   which is capable of conferring resistance to glyphosate.  The plasmid
CC   also contained the 1-aminocyclo-propane-1-carboxylic acid (ACC)
CC   deaminase gene from Pseudomonas chloroaphis (see also AAQ27199).  The
CC   transformed plants had ethylene levels reduced to about one half that
CC   of the control, untransformed plants.  It is expected that such plants
CC   will show reduced senesence of flowers and leaves when compared to
CC   untransformed plants.
XX
SQ   Sequence   455 AA;

  Query Match             100.0%;  Score 2288;  DB 13;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-181;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   2
AAW39426
ID   AAW39426 standard; Protein; 455 AA.
XX
AC   AAW39426;
XX
DT   19-MAY-1998  (first entry)
XX
DE   CP4-EPSPS protein.
XX
KW   1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;
KW   fruit ripening; ethylene production; glyphosate resistance;
KW   5-enolpyruvyl-3-shikimate phosphate synthase; EPSPS; CP4.
XX
OS   Synthetic.
XX
PN   US5702933-A.
XX
PD   30-DEC-1997.
XX
PF   06-NOV-1995;   95US-0553943.
XX
PR   17-DEC-1991;   91US-0809457.
PR   26-DEC-1990;   90US-0632440.
PR   06-NOV-1995;   95US-0553943.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Kishore GM,  Klee HJ;
XX
DR   WPI; 1998-076419/07.
DR   N-PSDB; AAV09719.
XX
PT   Production of plants with delayed ripening - using DNA encoding
PT   1-amino:cyclo:propane-1-carboxylic acid deaminase
XX
PS   Example 9; Fig 21; 56pp; English.
XX
CC   This sequence represents the synthetic CP4 5-enolpyruvyl-3-shikimate
CC   phosphate synthase (EPSPS) gene which is capable of conferring
CC   resistance to glyphosate and is used in a novel method for producing
CC   fruit-bearing plants with delayed ripening. The method involves the
CC   expression of a 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase
CC   gene in a plant at a level sufficient to reduce ethylene production in
CC   the fruit.
XX
SQ   Sequence   455 AA;

  Query Match             100.0%;  Score 2288;  DB 19;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-181;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   3
AAM52214
ID   AAM52214 standard; Protein; 527 AA.
XX
```

AC AAM52214;
XX
DT 12-FEB-2002 (first entry)
XX
DE EPSPS SEQ ID NO 3.
XX
KW Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;
KW 5-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;
KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;
KW chloroplast transit peptide.
XX
OS Petunia hybrida.
OS Agrobacterium sp.
XX
PN JP2001190168-A.
XX
PD 17-JUL-2001.
XX
PF 27-OCT-2000; 2000JP-0328811.
XX
PR 29-OCT-1999; 99JP-0310244.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2001-605307/69.
DR N-PSDB; ABA02854.
XX
PT New herbicide-resistant plant -
XX
PS Examples; Page 46-48; 66pp; Japanese.
XX
CC The invention relates to a transgenic plant which shows resistance to a
CC herbicide in an amount inhibiting natural
CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant
CC and having at least one enzymatic activity selected from:
CC (1) EPSPS activity different from natural EPSPS activity of the plant or
CC (2) glyphosate oxidoreductase (GOX) activity different from the natural
CC GOX activity of the plant; and in which a gene encoding a protein having
CC the following properties: (a) combines specifically with a substance
CC participating to the herbicidic activity a herbicide of
CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no
CC denaturing activity on a substance to which said protein combines
CC specifically; and (c) contains substantially no framework region of the
CC variable region of immunoglobulin. The present sequence is that of the
CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium
CC sp. strain CP4 EPSPS gene.
XX
SQ Sequence 527 AA;

Query Match 100.0%; Score 2288; DB 22; Length 527;
Best Local Similarity 100.0%; Pred. No. 4.2e-181;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 132

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     133 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 192

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     193 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 252

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 312

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     313 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 372

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 432

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     433 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 492

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||||
Db     493 vtvddatmiatsfpefmdlmaglgakielsdtkaa 527
```

RESULT 4
AAR22300
ID AAR22300 standard; Protein; 455 AA.
XX
AC AAR22300;
XX
DT 03-AUG-1992 (first entry)
XX
DE Class II EPSPS enzyme.
XX
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX
OS Agrobacterium sp. strain CP4.
XX
PN WO9204449-A.
XX
PD 19-MAR-1992.
XX
PF 28-AUG-1991; 91WO-US06148.
XX
PR 31-AUG-1990; 90US-0576537.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 3; 148pp; English.
XX
CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC phosphate synthase enzyme (EPSPS) of Agrobacterium sp. strain CP4
CC It is used to create glyphosate resistant plants or seeds which
CC can be planted in a field of crops to selectively control weeds.
CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC tobacco and alfalfa. This provides a cost effective, environmentally
CC compatible weed control device. See also AAR22301 and AAR22302.
XX
SQ Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 13; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.1e-180;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
                |||||||||||||||||||||||||||||||||
Db    421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

```
RESULT   5
AAW34683
ID   AAW34683 standard; Protein; 455 AA.
XX
AC   AAW34683;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Agrobacterium sp. strain CP4.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Agrobacterium sp. strain CP4.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          271..274
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93788.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Claim 7; Fig 3; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence is from
CC   Agrobacterium sp. strain CP4. The EPSPS enzymes are part of the shikimic
CC   acid pathway, which leads to the biosynthesis of aromatic compounds.
CC   EPSPS converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid
CC   to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the
CC   herbicide glyphosate. It would be useful to produce transgenic crops
CC   containing glyphosate resistance genes so that glyphosate-containing
CC   herbicides can be applied to selectively kill weeds. The novel EPSPS
CC   enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that
CC   when introduced into a plant, the plant is made glyphosate-tolerant, and
CC   EPSPS enzyme activity is not affected. These class II EPSPS enzymes are
CC   fused to a chloroplast transit peptide to target the protein into the
CC   chloroplast, which is the site for the shikimic acid pathway. In
CC   addition, the EPSPS gene is cloned into a plant under the control of a
CC   promoter such as figwort mosaic virus promoter or the cauliflower mosaic
CC   virus promoter, so that expression is enhanced.
XX
SQ   Sequence   455 AA;
```

```
  Query Match             99.7%;  Score 2282;  DB 18;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-180;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||||||||||||||||||||
Db    421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

```
RESULT   6
AAW24474
ID   AAW24474 standard; Protein; 455 AA.
XX
AC   AAW24474;
XX
DT   02-OCT-1997  (first entry)
XX
```

DE Class II EPSPS for glyphosate resistant plant production.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW tobacco.
XX
OS Agrobacterium sp. strain CP4.
XX
PN US5633435-A.
XX
PD 27-MAY-1997.
XX
PF 31-AUG-1990; 90US-0576537.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1997-297418/27.
DR N-PSDB; AAT77313.
XX
PT New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT used for transforming plants to produce plants which are tolerant to
PT glyphosate herbicide
XX
PS Claim 1; Column 57-60; 154pp; English.
XX
CC AAW24474 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC the production of herbicide resistant (glyphosate-tolerant) plants
CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 18; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.1e-180;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

RESULT 7
AAW71609
ID AAW71609 standard; Protein; 455 AA.
XX
AC AAW71609;
XX
DT 14-DEC-1998 (first entry)
XX
DE Agrobacterium sp. strain CP4 Class II EPSPS.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
KW N-phosphonomethylglycine.
XX
OS Agrobacterium sp.
XX
PN US5804425-A.
XX
PD 08-SEP-1998.
XX
PF 07-APR-1997; 97US-0833485.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
PR 07-APR-1997; 97US-0833485.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1998-505657/43.
DR N-PSDB; AAV58009.
XX
PT Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -
PT useful for characterisation of the enzyme to determine inhibition
PT data values
XX
PS Claim 3; Fig 3; 152pp; English.
XX
CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
CC produced by recombinant methods, can be used in kinetic studies to
CC determine Ki and Km values of the enzyme for its characterisation. The
CC enzyme is normally used for the production of 5-enolpyruvyl-3-
CC phosphoshikimic acid in plants, and most forms of the enzyme are
CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
CC Inhibition data enables more accurate values of concentrations of
CC herbicide to be used when growing the plant without being detrimental to
CC it. This enables the plant to be grown in the presence of the herbicide,
CC being used to inhibit the growth of undesired plants. The present
CC sequence represents a Class II EPSPS from bacterial isolate
CC Agrobacterium sp. strain CP4.
XX
SQ Sequence 455 AA;

Query Match 99.7%; Score 2282; DB 19; Length 455;
Best Local Similarity 99.8%; Pred. No. 1.1e-180;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
```

```
                 | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   8
AAE05053
ID   AAE05053 standard; Protein; 455 AA.
XX
AC   AAE05053;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Agrobacterium sp. strain CP4 class II EPSPS protein.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   transformed bacteria; transgenic plant; herbicide.
XX
OS   Agrobacterium sp. CP4.
XX
PN   US6248876-B1.
XX
PD   19-JUN-2001.
XX
PF   20-AUG-1998;   98US-0137440.
XX
PR   13-SEP-1994;   94US-0306063.
PR   07-APR-1997;   97US-0833485.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 2001-407326/43.
DR   N-PSDB; AAD09754.
XX
PT   DNA probe capable of use in a polymerase chain reaction for identifying
PT   the presence of a target genomic DNA encoding a
PT   5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme -
XX
PS   Claim 3; Fig 3; 152pp; English.
XX
CC   The present invention relates to a DNA probe capable of use in a
CC   polymerase chain reaction for identifying the presence of a target
CC   genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)
CC   enzyme referred to as class II EPSPS enzyme which is tolerant to
CC   glyphosate. EPSPS genes are useful in producing transformed bacteria
CC   and transgenic plants which are tolerant to glyphosate herbicide. The
CC   probe is useful for identifying the presence of a target genomic DNA
CC   encoding a EPSPS enzyme. The present sequence is Agrobacterium sp.
CC   strain CP4 class II EPSPS protein.
CC   Note: The present sequence, SED ID NO: 3 is stated as amino acid
CC   sequence throughout the specification. However, it is referred as probe
CC   in claim 3 of the specification.
XX
SQ   Sequence   455 AA;


  Query Match             99.7%;  Score 2282;  DB 22;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 1.1e-180;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   9
AAR22301
ID   AAR22301 standard; Protein; 449 AA.
XX
AC   AAR22301;
XX
DT   03-AUG-1992  (first entry)
XX
DE   Class II EPSPS enzyme.
XX
KW   Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW   5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX
OS   Achromobacter sp. strain LBAA.
XX
PN   WO9204449-A.
XX
PD   19-MAR-1992.
XX
PF   28-AUG-1991;   91WO-US06148.
XX
PR   31-AUG-1990;   90US-0576537.
```

XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 5; 148pp; English.
XX
CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC phosphate synthase enzyme (EPSPS) of Achromobacter sp. strain LBAA
CC It is used to create glyphosate resistant plants or seeds which
CC can be planted in a field of crops to selectively control weeds.
CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC tobacco and alfalfa. This provides a cost effective, environmentally
CC compatible weed control device. See also AAR22300 and AAR22302.
XX
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 10
AAR22302
ID AAR22302 standard; Protein; 449 AA.
XX
AC AAR22302;
XX
DT 03-AUG-1992 (first entry)
XX
DE Class II EPSPS enzyme.
XX
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.

XX
OS Pseudomonas sp. strain PG2982.
XX
PN WO9204449-A.
XX
PD 19-MAR-1992.
XX
PF 28-AUG-1991; 91WO-US06148.
XX
PR 31-AUG-1990; 90US-0576537.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 7; 148pp; English.
XX
CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC phosphate synthase enzyme (EPSPS) of Pseudomonas sp. strain PG2982
CC It is used to create glyphosate resistant plants or seeds which
CC can be planted in a field of crops to selectively control weeds.
CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC tobacco and alfalfa. This provides a cost effective, environmentally
CC compatible weed control device. See also AAR22300 and AAR22301.
XX
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 13; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 11
AAW34684

```
ID   AAW34684 standard; Protein; 449 AA.
XX
AC   AAW34684;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Achromobacter sp. strain LBAA.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Achromobacter sp. strain LBAA.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          271..274
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93789.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Disclosure; Fig 4; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence was isolated from
CC   Achromobacter sp. strain LBAA. The EPSPS enzymes are part of the
CC   shikimic acid pathway, which leads to the biosynthesis of aromatic
CC   compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and
CC   3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC   inhibited by the herbicide glyphosate. It would be useful to produce
CC   transgenic crops containing glyphosate resistance genes so that
CC   glyphosate-containing herbicides can be applied to selectively kill
CC   weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for
CC   glyphosate, such that when introduced into a plant, the plant is made
CC   glyphosate-tolerant, and EPSPS enzyme activity is not affected. These
CC   class II EPSPS enzymes are fused to a chloroplast transit peptide to
CC   target the protein into the chloroplast, which is the site for the
CC   shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant
CC   under the control of a promoter such as figwort mosaic virus promoter or
CC   the cauliflower mosaic virus promoter, so that expression is enhanced.
XX
SQ   Sequence   449 AA;
```

```
  Query Match             83.1%;  Score 1900.5;  DB 18;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 4.3e-149;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| |||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :|||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447


RESULT  12
AAW34685
ID   AAW34685 standard; Protein; 449 AA.
XX
AC   AAW34685;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Pseudomonas sp. strain PG2982.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Pseudomonas sp. strain PG2982.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          273..276
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
```

```
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93790.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Disclosure; Fig 5; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence was isolated from
CC   Pseudomonas sp. strain PG2982. The EPSPS enzymes are part of the
CC   shikimic acid pathway, which leads to the biosynthesis of aromatic
CC   compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and
CC   3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC   inhibited by the herbicide glyphosate. It would be useful to produce
CC   transgenic crops containing glyphosate resistance genes so that
CC   glyphosate-containing herbicides can be applied to selectively kill
CC   weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for
CC   glyphosate, such that when introduced into a plant, the plant is made
CC   glyphosate-tolerant, and EPSPS enzyme activity is not affected. These
CC   class II EPSPS enzymes are fused to a chloroplast transit peptide to
CC   target the protein into the chloroplast, which is the site for the
CC   shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant
CC   under the control of a promoter such as figwort mosaic virus promoter or
CC   the cauliflower mosaic virus promoter, so that expression is enhanced.
XX
SQ   Sequence   449 AA;

  Query Match             83.1%;  Score 1900.5;  DB 18;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 4.3e-149;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||||  || ||| ||||||  || ||||||||  ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | |||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447


RESULT  13
AAW24479
ID   AAW24479 standard; Protein; 449 AA.
XX
AC   AAW24479;
XX
DT   02-OCT-1997  (first entry)
XX
DE   Class II EPSPS used for glyphosate resistant plant production.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW   glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW   fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW   tobacco.
XX
OS   Achromobacter sp. strain LBAA.
XX
PN   US5633435-A.
XX
PD   27-MAY-1997.
XX
PF   31-AUG-1990;   90US-0576537.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-297418/27.
DR   N-PSDB; AAT77315.
XX
PT   New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT   used for transforming plants to produce plants which are tolerant to
PT   glyphosate herbicide
XX
PS   Claim 7; Column 63-66; 154pp; English.
XX
CC   AAW24479 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC   phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC   to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC   the production of herbicide resistant (glyphosate-tolerant) plants
CC   such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC   rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC   pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ   Sequence   449 AA;

  Query Match             83.1%;  Score 1900.5;  DB 18;  Length 449;
  Best Local Similarity   82.9%;  Pred. No. 4.3e-149;
  Matches  373;  Conservative   32;  Mismatches   42;  Indels    3;  Gaps    1;

Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||||  || ||| ||||||  || ||||||||  ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| ||||  ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180
```

```
Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db    241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db    301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db    361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: |||||||||||:| |||||||||
Db    418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 14
AAW24480
ID AAW24480 standard; Protein; 449 AA.
XX
AC AAW24480;
XX
DT 02-OCT-1997 (first entry)
XX
DE Class II EPSPS used for glyphosate resistant plant production.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW tobacco.
XX
OS Pseudomonas sp. strain PG2982.
XX
PN US5633435-A.
XX
PD 27-MAY-1997.
XX
PF 31-AUG-1990; 90US-0576537.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1997-297418/27.
DR N-PSDB; AAT77316.
XX
PT New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT used for transforming plants to produce plants which are tolerant to
PT glyphosate herbicide
XX
PS Claim 7; Column 69-72; 154pp; English.
XX
CC AAW24480 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC the production of herbicide resistant (glyphosate-tolerant) plants
CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 18; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

```
Qy      1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          | | || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||
Db      1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          :|||||||:|||||| |||:|||||| || ||| |||||| || |||||||| ||  ::||
Db     61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||:|||||||||||||||:: |||||:|:|| |||| ||||||||||||||||||
Db    121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||| |||||| |||||||||||||||||||||||||||||:| |||||||
Db    241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db    301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db    361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: |||||||||||:| |||||||||
Db    418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 15
AAW71611
ID AAW71611 standard; Protein; 449 AA.
XX
AC AAW71611;
XX
DT 14-DEC-1998 (first entry)
XX
DE Pseudomonas sp. strain PG2982 Class II EPSPS.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
KW N-phosphonomethylglycine.
XX
OS Pseudomonas sp.
XX
PN US5804425-A.
XX
PD 08-SEP-1998.
XX
PF 07-APR-1997; 97US-0833485.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
PR 07-APR-1997; 97US-0833485.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1998-505657/43.
DR N-PSDB; AAV58012.
XX
PT Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -

PT useful for characterisation of the enzyme to determine inhibition
PT data values
XX
PS Disclosure; Fig 5; 152pp; English.
XX
CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
CC produced by recombinant methods, can be used in kinetic studies to
CC determine Ki and Km values of the enzyme for its characterisation. The
CC enzyme is normally used for the production of 5-enolpyruvyl-3-
CC phosphoshikimic acid in plants, and most forms of the enzyme are
CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
CC Inhibition data enables more accurate values of concentrations of
CC herbicide to be used when growing the plant without being detrimental to
CC it. This enables the plant to be grown in the presence of the herbicide,
CC being used to inhibit the growth of undesired plants. The present
CC sequence represents a Class II EPSPS from bacterial isolate
CC Pseudomonas sp. strain PG2982.
XX
SQ Sequence 449 AA;

Query Match 83.1%; Score 1900.5; DB 19; Length 449;
Best Local Similarity 82.9%; Pred. No. 4.3e-149;
Matches 373; Conservative 32; Mismatches 42; Indels 3; Gaps 1;

```
Qy       1 MLHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | | || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   |  |||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

Search completed: August 15, 2002, 13:58:48
Job time: 67 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:59:21 ; Search time 17.59 Seconds
(without alignments)
1001.557 Million cell updates/sec

Title: US-09-464-099A-3
Perfect score: 2288
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2288	100.0	455	1	AROA_AGRSP	Q9r4e4 agrobacteri
2	1906.5	83.3	449	1	AROA_PSES2	P56952 pseudomonas
3	884	38.6	431	1	AROA_BACHD	Q9kca6 bacillus ha
4	867.5	37.9	447	1	AROA_SYNY3	Q59975 synechocyst
5	806.5	35.2	430	1	AROA_LACLA	Q9ceu0 lactococcus
6	806	35.2	443	1	AROA_BACNO	Q46550 bacteroides
7	803.5	35.1	427	1	AROA_STRPN	Q9s400 streptococc
8	802.5	35.1	430	1	AROA_LACLC	P43905 lactococcus
9	785	34.3	428	1	AROA_BACSU	P20691 bacillus su
10	766.5	33.5	431	1	AROA_AQUAE	O67494 aquifex aeo
11	695	30.4	410	1	AROA_THEMA	Q9wyi0 thermotoga
12	590	25.8	428	1	AROA_CAMJE	P52312 campylobact
13	577	25.2	430	1	AROA_STAAU	Q05615 staphylococ
14	521.5	22.8	429	1	AROA_HELPJ	Q9zkf7 helicobacte
15	518.5	22.7	429	1	AROA_HELPY	P56197 helicobacte
16	496	21.7	419	1	AROA_METTH	O26860 methanobact
17	459.5	20.1	429	1	AROA_METJA	Q57925 methanococc
18	408	17.8	439	1	AROA_DEIRA	Q9rvd3 deinococcus
19	403	17.6	416	1	AROA_ARCFU	O28775 archaeoglob
20	402.5	17.6	427	1	AROA_AERPE	Q9yek9 aeropyrum p
21	402.5	17.6	427	1	AROA_YEREN	P19688 yersinia en
22	395	17.3	428	1	AROA_YERPE	Q60112 yersinia pe
23	387.5	16.9	432	1	AROA_HAESO	P52310 haemophilus
24	378	16.5	427	1	AROA_ECOLI	P07638 escherichia
25	376	16.4	427	1	AROA_SHISO	Q9zff7 shigella so
26	373.5	16.3	410	1	AROA_PYRAB	Q9v1h1 pyrococcus
27	373	16.3	432	1	AROA_PASHA	P54220 pasteurella
28	372	16.3	427	1	AROA_KLEPN	P24497 klebsiella
29	372	16.3	440	1	AROA_PASMU	Q04570 pasteurella
30	371	16.2	427	1	AROA_SHIDY	O87006 shigella dy
31	365.5	16.0	463	1	AROA_BURPS	P39915 burkholderi
32	365	16.0	428	1	AROA_EDWIC	Q9x4h2 edwardsiell
33	363.5	15.9	432	1	AROA_HAEIN	Q03421 haemophilus
34	357	15.6	427	1	AROA_SALTY	P07637 salmonella
35	348	15.2	427	1	AROA_SALGL	P22299 salmonella
36	347	15.2	427	1	AROA_SALTI	P19786 salmonella
37	323.5	14.1	516	1	AROA_PETHY	P11043 petunia hyb
38	321.5	14.1	1588	1	ARO1_YEAST	P08566 s pentafunc
39	317	13.9	441	1	AROA_CHLMU	Q9pk28 chlamydia m
40	315.5	13.8	518	1	ARO1_TOBAC	P23981 nicotiana t
41	314.5	13.7	440	1	AROA_CHLTR	O84371 chlamydia t
42	313	13.7	428	1	AROA_BUCAP	Q59178 buchnera ap
43	310.5	13.6	520	1	AROA_LYCES	P10748 lycopersico
44	308.5	13.5	1573	1	ARO1_SCHPO	Q9p7r0 s pentafunc
45	306.5	13.4	445	1	AROA_CHLPN	Q9z6m0 chlamydia p

ALIGNMENTS

RESULT 1
AROA_AGRSP
ID AROA_AGRSP STANDARD; PRT; 455 AA.
AC Q9R4E4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN AROA.
OS Agrobacterium sp. (strain CP4).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=361;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.
RA Barry G.F., Kishore G.M., Padgette S.R., Stallings W.C.;
RT "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL Patent number US5633435, 27-MAY-1997.
RN [2]
RP SEQUENCE OF 1-15.
RX MEDLINE=96182485; PubMed=8598558;
RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,
RA Nida D.L., Burnette B.L., Nickson T.E., Mitsky T.A., Taylor M.L.,
RA Fuchs R.L., Padgette S.R.;
RT "The expressed protein in glyphosate-tolerant soybean, 5-
RT enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp.
RT strain CP4, is rapidly digested in vitro and is not toxic to acutely
RT gavaged mice.";
RL J. Nutr. 126:728-740(1996).
CC -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in
CC glyphosate-tolerant soybean, canola, cotton and maize by Monsanto.
CC Developed to provide new weed-control options for farmers.
CC Expression of this protein in plants imparts high levels of
CC glyphosate tolerance.
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR InterPro; IPR001986; EPSP_syntase.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance;
KW Genetically modified food.
FT CONFLICT 2 2 S -> L (IN REF. 2).
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;

Query Match 100.0%; Score 2288; DB 1; Length 455;
Best Local Similarity 100.0%; Pred. No. 3.8e-135;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
```

```
RESULT   2
AROA_PSES2
ID   AROA_PSES2     STANDARD;      PRT;   449 AA.
AC   P56952;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Pseudomonas sp. (strain PG2982), and
OS   Achromobacter sp. (strain LBAA).
OC   Bacteria; Proteobacteria.
OX   NCBI_TaxID=308, 129026;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 2-16.
RA   Barry G.F., Kishore G.M., Padgette S.R., Stallings W.C.;
RT   "Glyphosate-tolerant 5-enolpyruvylshikimate-3-phosphate synthases.";
RL   Patent number US5633435, 27-MAY-1997.
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- MISCELLANEOUS: RESISTANT TO THE ANTIBIOTIC GLYPHOSATE.
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Herbicide resistance.
SQ   SEQUENCE   449 AA;  47297 MW;  447F213EECCAEFC1 CRC64;
```

```
  Query Match           83.3%;  Score 1906.5;  DB 1;  Length 449;
  Best Local Similarity  83.1%;  Pred. No. 1.8e-111;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;
```

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 MSHSASPKPATARRSEALTGEIRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||||  || ||| ||||||  || ||||||||  ||  ::||
Db      61 RAMQAMGAKIRKEGDVWIINGVGNGCLLQPEAALDFGNAGTGARLTMGLVGTYDMKTSFI 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||:: ||||:|:|| ||||  ||||||||||||||||
Db     121 GDASLSKRPMGRVLNPLREMGVQVEAADGDRMPLTLIGPKTANPITYRVPMASAQVKSAV 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRITGQGKLVGQTID 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db     241 VPGDPSSTAFPLVAALLVEGSDVTIRNVLMNPTRTGLILTLQEMGADIEVLNARLAGGED 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 VADLRVRASKLKGVVVPPERAPSMIDEYPVLAIAASFAEGETVMDGLDELRVKESDRLAA 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   | |||||||||||||||||| :| |
Db     361 VARGLEANGVDCTEGEMSLTVRGRPDGKGLG---GGTVATHLDHRIAMSFLVMGLAAEKP 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 VTVDDSNMIATSFPEFMDMMPGLGAKIELS 447
```

```
RESULT   3
AROA_BACHD
ID   AROA_BACHD     STANDARD;      PRT;   431 AA.
AC   Q9KCA6;
DT   01-MAR-2002 (Rel. 41, Created)
DT   01-MAR-2002 (Rel. 41, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR AROE OR BH1667.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN THE
CC       BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; AP001512; BAB05386.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   431 AA;  45485 MW;  12F4FFBE7BA0743D CRC64;


  Query Match          38.6%;  Score 884;  DB 1;  Length 431;
  Best Local Similarity   44.2%;  Pred. No. 4.7e-48;
  Matches  188;  Conservative   72;  Mismatches  153;  Indels   12;  Gaps    5;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           : || ||:::|||||||||: ||| || | | : | | | ::|    | :|  | :
Db      10 AKGLKGTIKVPGDKSISHRAVMFGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |  ||  |   || ||: |  || :|::    | |  ||| |: |||| ||
Db      70 ERVTVKGKGWDGLREPSDILDVGNSGTTTRLILGILSTLPFHSVIIGDESIGKRPMKRVT 129

Qy     135 NPLREMGVQVKSED-GDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
            ||: || |:   | |:  |:::|| :    | :  |:||||:|||:|||||   | |:|
Db     130 EPLKSMGAQIDGRDHGNLTPLSIRGGQL-KGIDFHSPVASAQMKSAILLAGLRAEGKTSV 188

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
            ||  ||||||:||: || |:    : ||: |: :||   |||| : |||| || || ||
Db     189 TEPAKTRDHTERMLEAFGVNI----EKDGL-TVSIEGGQMLTGQHVVVPGDISSAAFFLV 243

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
           |  :|| | :|: || :|||| |::  |::||| : : | |: ||| |||| : :| |:|
Db     244 AGAMVPHSRITLTNVGINPTRAGILEVLKQMGATLAMENERVQGGEPVADLTIETSVLQG 303

Qy     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373
           | :  |  | :||| ||:|| |  | | ||:   |||:|||::|:  | : |   |
Db     304 VEIGGDIIPRLIDEIPIIAVLATQASGRTVIKDAEELKVKETNRIDTVVSELTKLGASIH 363

Qy     374 EGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSF 433
           :  ::: |    ||     |  |::| |||| |:  :  |::| ||||:    || |:
Db     364 ATDDGMIIEGPTPLKG-----GVTVSSHGDHRIGMAMAIAALLAEKPVTVEGTEAIAVSY 418

Qy     434 PEFMD 438
           | | |
Db     419 PSFFD 423


ULT   4
OA_SYNY3
ID   AROA_SYNY3     STANDARD;      PRT;   447 AA.
AC   Q59975; Q59974;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR SLR0444.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94299161; PubMed=8026753;
RA   dalla Chiesa M., Mayes S.R., Maskell D.S., Nixon P.J., Barber J.;
RT   "An aroA homologue from Synechocystis sp. PCC 6803.";
RL   Gene 144:145-146(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97061201; PubMed=8905231;
RA   Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA   Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA   Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA   Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA   Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT   entire genome and assignment of potential protein-coding regions.";
RL   DNA Res. 3:109-136(1996).
RN   [3]
RP   SEQUENCE OF 103-194 FROM N.A.
RX   MEDLINE=93307506; PubMed=7686511;
RA   Mayes S.R., dalla Chiesa M., Zhang Z., Barber J.;
RT   "The genes aroA and trnQ are located upstream of psbO in the
RT   chromosome of Synechocystis 6803.";
RL   FEBS Lett. 325:255-261(1993).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X75325; CAA53074.1; -.
DR   EMBL; D90914; BAA18477.1; -.
DR   EMBL; X72784; CAA51291.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   447 AA;  47046 MW;  2F3C8AD26B5A7BCE CRC64;


  Query Match          37.9%;  Score 867.5;  DB 1;  Length 447;
  Best Local Similarity   48.1%;  Pred. No. 5.2e-47;
  Matches  203;  Conservative   62;  Mismatches  144;  Indels   13;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRK-EGDT 76
           |:| :|:|||||||||: | | :|:||| | ||| |||  :|    :|||| | :   :
Db      26 LTGRLRVPGDKSISHRALMLGAIATGETIIEGLLLGEDPRSTAHCFRAMGAEISELNSEK 85

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLV-GVYDFDSTFIGDASLTKRPMGRVLN 135
           |: | | | |   || ||: |  || :||: |  |   |  || ||  ||| ||:
Db      86 IIVQGRGLGQLQEPSTVLDAGNSGTTMRLMLGLLAGQKDCLFTVTGDDSLRHRPMSRVIQ 145

Qy     136 PLREMGVQV-KSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
           ||::|| ::    :|   |: ::| :   || |  |:|||||| :|||| | | |||
Db     146 PLQQMGAKIWARSNGKFAPLAVQGSQL-KPIHYHSPIASAQVKSCLLLAGLTTEGDTTVT 204

Qy     195 EPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVA 254
           || ::|||:|:||| ||| ||::       :: : |   |||| : |||| || || |||
Db     205 EPALSRDHSERMLQAFGAKLTID---PVTHSVTVHGPAHLTGQRVVVPGDISSAAFWLVA 261

Qy     255 ALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGV 314
           | ::|||:: : || :||||||::  | :|||||   | ||  || |||||||:| |:|
Db     262 ASILPGSELLVENVGINPTRTGVLEVLAQMGADITPENERLVTGEPVADLRVRASHLQGC 321

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
           |   :  | :||| |||||||||||| | :    ||||||||||:|:|: |   |    |
Db     322 TFGGEIIPRLIDEIPILAVAAAFAEGTTRIEDAAELRVKESDRLAAIASELGKMGAKVTE 381

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFP 434
           :  | ::|   | |    || | :  ||||||:  :  | |     :: |   | |:|
Db     382 FDDGLEIQG---GSPL---QGAEVDSLTDHRIAMALAIAALGSGGQTIINRAEAAAISYP 435

Qy     435 EF 436
```

```
Db    436 EF 437


RESULT   5
AROA_LACLA
ID   AROA_LACLA     STANDARD;      PRT;   430 AA.
AC   Q9CEU0;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR LL1744.
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IL1403;
RX   MEDLINE=21235186; PubMed=11337471;
RA   Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT   "The complete genome sequence of the lactic acid bacterium Lactococcus
RT   lactis ssp. lactis IL1403.";
RL   Genome Res. 11:731-753(2001).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE006404; AAK05842.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   430 AA;  45958 MW;  CB216F07AA4EE799 CRC64;


  Query Match             35.2%;  Score 806.5;  DB 1;  Length 430;
  Best Local Similarity   41.9%;  Pred. No. 3e-43;
  Matches  184;  Conservative   77;  Mismatches  157;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| |||::| :| :|||  |  :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTVIYDILRGEDVLSTIEAFRAMGVEIEDKG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
             :  : | |   | |||  || ||: |  ||  |::   |::|  || ||:|||| ||
Db      67 EVITVHGKGISELKAPEKALDMGNSGTSTRLLSGILAGLPFETTLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|      | | :|:|||||||||: | |   |:| |
Db     127 TPLQLMGAEITGQTDKVKLPMTIKGSTHLKAIDYVLPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           * :|  || |||:|| ||  | |       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGELKVSD-----KTILVPGGQKLVGQKVVVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           |||:|  |:: : || :| ||||:|  :| ||  :|::       :|     | |:|::|
Db     242 AALVVENSELILENVGVNETRTGIIEVIQAMGGQLEILE------QDNVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
           |||  :  |  | :||| ||:|: |  |:| |::    ||:|||:||:: ||| |   |
Db     296 QLKGTEISGDLIPRLIDELPIIALLATQAQGQTIIRDAAELKVKETDRIAVVANALNSMG 355

Qy     370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428
             :  ::::|   |  | :|   :: |  |||| |   :  |: :|  : :: |
Db     356 AKIQPTDDGMIIQG---GTKL-HAPENSINTLGDHRIGMMAAIAALLVKNGEIELERAEA 411

Qy     429 IATSFPEFMDLMAGLGAKI 447
           | ||:| | | :  |   |
Db     412 IQTSYPSFFDDLEQLSENI 430


RESULT   6
AROA_BACNO
ID   AROA_BACNO     STANDARD;      PRT;   443 AA.
AC   Q46550;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Bacteroides nodosus (Dichelobacter nodosus).
OC   Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OC   Dichelobacter.
OX   NCBI_TaxID=870;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VCS1001;
RX   MEDLINE=94320795; PubMed=8045432;
RA   Alm R.A., Dalrymple B.P., Mattick J.S.;
RT   "Sequencing and expression of the aroA gene from Dichelobacter
RT   nodosus.";
RL   Gene 145:97-101(1994).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z29339; CAA82544.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   443 AA;  47444 MW;  BE2243277ADEFBD5 CRC64;


  Query Match             35.2%;  Score 806;  DB 1;  Length 443;
  Best Local Similarity   42.2%;  Pred. No. 3.4e-43;
  Matches  185;  Conservative   74;  Mismatches  169;  Indels   10;  Gaps    3;

Qy      16 SGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75
           | ||| : | ||||:|||: :   || |:| | | |   | : | :|::|:|  |::| :
Db      12 SALSGEITICGDKSMSHRALLLAALAEGQTEIRGFLACADCLATRQALRALGVDIQREKE 71
```

```
Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db      72 IVTIRGVGFLGLQPPKAPLNMQNSGTSMRLLAGILAAQRFESVLCGDESLEKRPMQRIIT 131

Qy     136 PLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
Db     132 PLVQMGAKIVSHSNFTAPLHISG-RPLTGIDYALPLPSAQLKSCLILAGLLADGTTRLHT 190

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
Db     191 CGISRDHTERMLPLFGGALEIKKE-----QIIVTGGQKLHGCVLDIVGDLSAAAFFMVAA 245

Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
Db     246 LIAPRAEVVIRNVGINPTRAAIITLLQKMGGRIELHHQRFWGAEPVADIVVYHSKLRGIT 305

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
Db     306 VAPEWIANAIDELPIFFIAAACAEGTTFVGNLSELRVKESDRLAAMAQNLQTLGVACDVG 365

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPE 435
Db     366 ADFIHIYGRSDRQFL----PARVNSFGDHRIAMSLAVAGVRAAGELLIDDGAVAAVSMPQ 421

Qy     436 FMDLMAGLGAKIELSDTK 453
Db     422 FRDFAAAIGMNVGEKDAK 439
```

```
RESULT   7
AROA_STRPN
ID   AROA_STRPN     STANDARD;      PRT;   427 AA.
AC   Q9S400;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   01-MAR-2002 (Rel. 41, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR SP1371.
OS   Streptococcus pneumoniae.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1313;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=NCIMB 40794 / 0100993;
RX   MEDLINE=20069365; PubMed=10601870;
RA   Du W., Wallis N.G., Mazzulla M.J., Chalker A.F., Zhang L., Liu W.-S.,
RA   Kallender H., Payne D.J.;
RT   "Characterization of Streptococcus pneumoniae 5-enolpyruvylshikimate
RT   3-phosphate synthase and its activation by univalent cations.";
RL   Eur. J. Biochem. 267:222-227(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TIGR4;
RX   MEDLINE=21357209; PubMed=11463916;
RA   Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA   Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA   Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA   Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA   Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA   McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA   Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA   Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT   "Complete genome sequence of a virulent isolate of Streptococcus
RT   pneumoniae.";
RL   Science 293:498-506(2001).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF169483; AAD45819.1; -.
DR   EMBL; AE007434; AAK75469.1; -.
DR   TIGR; SP1371; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT   CONFLICT     13     13       S -> I (IN REF. 1).
FT   CONFLICT     71     71       I -> V (IN REF. 1).
FT   CONFLICT    201    201       K -> Q (IN REF. 1).
FT   CONFLICT    303    303       G -> C (IN REF. 1).
SQ   SEQUENCE   427 AA;  45766 MW;  45CE6F4D0D1C7B70 CRC64;
```

```
  Query Match             35.1%;  Score 803.5;  DB 1;  Length 427;
  Best Local Similarity   42.8%;  Pred. No. 4.6e-43;
  Matches  184;  Conservative   72;  Mismatches  155;  Indels   19;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
Db      10 LHGSIRVPGDKSISHRSIIFGSLAEGETKVYDILRGEDVLSTMQVFRDLGVEIEDKDGVI 69

Qy      78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
Db      70 TIQGVGMAGLKAPQNALNMGNSGTSIRLISGVLAGADFEVEMFGDDSLSKRPMDRVTLPL 129

Qy     138 REMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196
Db     130 KKMGVSISGQTERDLPPLRLKGTKNLRPIHYELPIASAQVKSALMFAALQAKGESVIIEK 189

Qy     197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
Db     190 EYTRNHTEDMLKQFGGHLSV----DG-KKITVQGPQKLTGQKVVVPGDISSAAFWLVAGL 244

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEV--INPRLAGGEDVADLRVRSSTLKGV 314
Db     245 IAPNSRLVLQNVGINETRTGIIDVIRAMGGKLEITEIDPVAKS----ATLIVESSDLKGT 300

Qy     315 TVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDE 374
Db     301 EIGGALIPRLIDELPIIALLATQAQGVTVIKDAEELKVKETDRIQVVADALNSMGADITP 360

Qy     375 GETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSF 433
Db     361 TADGMIIKGK------SALHGARVNTFGDHRIGMMTAIAALLVADGEVELDRAEAINTSY 414

Qy     434 PEFMDLMAGL 443
Db     415 PSFFDDLESL 424
```

```
RESULT   8
AROA_LACLC
ID   AROA_LACLC     STANDARD;      PRT;   430 AA.
AC   P43905;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
```

```
OS   Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1359;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MG1363 / F15876;
RX   MEDLINE=95124293; PubMed=7823907;
RA   Griffin H.G., Gasson M.J.;
RT   "Genetic aspects of aromatic amino acid biosynthesis in Lactococcus
RT   lactis.";
RL   Mol. Gen. Genet. 246:119-127(1995).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78413; CAA55180.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   430 AA;  45804 MW;  C5A197A49072C9D5 CRC64;
```

```
  Query Match             35.1%;  Score 802.5;  DB 1;  Length 430;
  Best Local Similarity   41.2%;  Pred. No. 5.4e-43;
  Matches  181;  Conservative   79;  Mismatches  158;  Indels   21;  Gaps    7;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           | || | :::|||||||||| ||| :| |:| |  :| |||||::| :| :|:|  |  :|
Db       7 SQGLKGRLKVPGDKSISHRSIMFGSIAKGKTIIHDILRGEDVLSTIEAFRALGVEIEDDG 66

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
              : | |   |  ||  || ||: |  ||  |::    |::|  || ||:|||| ||
Db      67 QVITVHGQGISKLKEPEKALDMGNSGTSTRLLSGILAGLPFEATLFGDDSLSKRPMDRVA 126

Qy     135 NPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || ::  : |  :||:|::|     | | :|:||||||||: | |   |:| |
Db     127 TPLQMMGAEIVGQTDKVKLPMTIKGSAHLKAIDYILPVASAQVKSAVIFAALQAEGLTKV 186

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           :|   || |||:||  ||  :||       :|| : |  || || : |||| || || ||
Db     187 VEKEKTRSHTEEMLVQFGGEITVSD-----KTILVPGGQKLLGQEVTVPGDISSAAFWLV 241

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED----VADLRVRSS 309
           | |:|  | : : || :| ||||::  :| ||  :|::       :|     | |:|::|
Db     242 AGLVVENSGLILENVGINETRTGILEVIQAMGGQLEILE------QDEVAKAATLKVKAS 295

Qy     310 TLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNG 369
            |||  :  |  | :||| ||:|: |  ||| |::    ||:|||:||:: ||: |   |
Db     296 QLKGTEISGDLIPRLIDELPIIALLATQAEGKTIIRDAAELKVKETDRIAVVADALNSMG 355

Qy     370 VDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN-PVTVDDATM 428
           : :  :  ::::|   |  | :|   |: |  |||| |   :  |: ||  : :: |
Db     356 ANIEPTDDGMIIQG---GTKL-HAPENAINTLGDHRIGMMVAIAALLVENGEIELERAEA 411

Qy     429 IATSFPEFMDLMAGLGAKI 447
           | ||:| | | :  |   :
Db     412 IQTSYPSFFDDLEKLSGNL 430
```

```
RESULT   9
AROA_BACSU
ID   AROA_BACSU     STANDARD;      PRT;   428 AA.
AC   P20691;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROE.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87191994; PubMed=3106153;
RA   Henner D.J., Band L., Flaggs G., Chen E.;
RT   "The organization and nucleotide sequence of the Bacillus subtilis
RT   hisH, tyrA and aroE genes.";
RL   Gene 49:147-152(1986).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M80245; AAA20869.1; -.
DR   EMBL; Z99115; CAB14176.1; -.
DR   PIR; C26532; C26532.
DR   SubtiList; BG10294; aroE.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   428 AA;  45240 MW;  DE3F7B96E761CB40 CRC64;
```

```
  Query Match             34.3%;  Score 785;  DB 1;  Length 428;
  Best Local Similarity   41.0%;  Pred. No. 6.5e-42;
  Matches  178;  Conservative   68;  Mismatches  176;  Indels   12;  Gaps    4;

Qy      14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73
           |   | | : |||||||||| ||| ||:| | :   | | | ::|    : ||  | :
Db       5 KVQTLHGEIHIPGDKSISHRSVMFGALAAGTTTVKNFLPGADCLSTIDCFRKMGVHIEQS 64

Qy      74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133
              :| | |   |  ||: || ||: |  || :|::    | |   || |: |||| ||
Db      65 SSDVVIHGKGIDALKEPESLLDVGNSGTTIRLMLGILAGRPFYSAVAGDESIAKRPMKRV 124

Qy     134 LNPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192
             ||::|| ::     |:  |::: |      | |  |:||||:|||||||||   | ||
Db     125 TEPLKKMGAKIDGRAGGEFTPLSVSGASL-KGIDYVSPVASAQIKSAVLLAGLQAEGTTT 183

Qy     193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252
           | ||  :|||||:|| ||  |     ::   :: : |  |||   | |||| || || |
Db     184 VTEPHKSRDHTERMLSAFGVKL-----SEDQTSVSIAGGQKLTAADIFVPGDISSAAFFL 238

Qy     253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLK 312
```

```
Db    239 AAGAMVPNSRIVLKNVGLNPTRTGIIDVLQNMGAKLEIKPSADSGAEPYGDLIIETSSLK 298

Qy    313 GVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDC 372

Db    299 AVEIGGDIIPRLIDEIPIIALLATQAEGTTVIKDAAELKVKETNRIDTVVSELRKLGAEI 358

Qy    373 DEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATS 432

Db    359 EPTADGMKVYGKQTLKG-----GAAVSSHGDHRIGMMLGIASCITEEPIEIEHTDAIHVS 413

Qy    433 FPEFMDLMAGLGAK 446

Db    414 YPTFFEHLNKLSKK 427


RESULT  10
AROA_AQUAE
ID   AROA_AQUAE     STANDARD;      PRT;   431 AA.
AC   O67494;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR AQ_1536.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000744; AAC07443.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   431 AA;  47793 MW;  EF842512EBE41D2A CRC64;


  Query Match             33.5%;  Score 766.5;  DB 1;  Length 431;
  Best Local Similarity   40.7%;  Pred. No. 9.2e-41;
  Matches  174;  Conservative   73;  Mismatches  169;  Indels   11;  Gaps    6;

Qy     14 KSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKE 73

Db      6 KIKRVKGELRVPSDKSITHRAFILGALASGETLVRKPLISGDTLATLEILKAIRTKVREG 65

Qy     74 GDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRV 133

Db     66 KEEVLIEG-RNYTFLEPHDVLDAKNSGTTARIMSGVLSTQPFFSVLTGDESLKNRPMLRV 124

Qy    134 LNPLREMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITT 192

Db    125 VEPLREMGAKIDGREEGNKLPIAIRGGNL-KGISYFNKKSSAQVKSALLLAGLRAEGMTE 183

Qy    193 VIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPL 252

Db    184 VVEPYLSRDHTERMLKLFGAEVITIPEERG-HIVKIKGGQELQGTEVYCPADPSSAAYFA 242

Qy    253 VAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRS-STL 311

Db    243 ALATLAPEGEIRLKEVLLNPTRDGFYRKLIEMGGDISFENYRELSNEPMADLVVRPVDNL 302

Qy    312 KGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVD 371

Db    303 KPVKVSPEEVPTLIDEIPILAVLMAFADGVSEVKGAKELRYKESDRIKAIVTNLRKLGVQ 362

Qy    372 CDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIAT 431

Db    363 VEEFEDGFAIHGTKEIK------GGVIETFKDHRIAMAFAVLGLVVEEEVIIDHPECVTV 416

Qy    432 SFPEFMD 438

Db    417 SYPEFWE 423


RESULT  11
AROA_THEMA
ID   AROA_THEMA     STANDARD;      PRT;   410 AA.
AC   Q9WYI0;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR TM0345.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogales; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001715; AAD35431.1; -.
DR   TIGR; TM0345; -.
```

```
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   410 AA;  45341 MW;  DC9F7D44792CA69F CRC64;


  Query Match             30.4%;  Score 695;  DB 1;  Length 410;
  Best Local Similarity   40.5%;  Pred. No. 2.4e-36;
  Matches  170;  Conservative   67;  Mismatches  161;  Indels   22;  Gaps   10;

Qy      24 IPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVG 83
           :| ||||:||: :   ||  |: :  ||   |   |   :: :| |   ||| |    |
Db       4 VPPDKSITHRALILSALAETESTLYNLLRCLDTERTHDILEKLGTRF--EGD-WEKMKVF 60

Qy      84 NGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLREMGVQ 143
                |  ||  ||: |  ||  |:  |:  :   || ||::||| ||: ||  || :
Db      61 PKPFAEPIEPLFCGNSGTTTRLMSGVLASYEMFTVLYGDPSLSRRPMRRVIEPLEMMGAR 120

Qy     144 VKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIMTRDHT 203
             :   : ||: ::|    : |:|: |:||||||||||||||   | | |||| :|||
Db     121 FMARQNNYLPMAIKGNHL-SGISYKTPVASAQVKSAVLLAGLRASGRTIVIEPAKSRDHT 179

Qy     204 EKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVPGSDV 263
           |:||:    || |  : :| | : ||      |  : |||| || || :|   : | : :
Db     180 ERMLK----NLGVPVEVEGTRVV-LE-PATFRGFTMKVPGDISSAAFFVVLGAIHPNARI 233

Qy     264 TILNVLMNPTRTGLILTLQEMGADI--EVINPRLAGGEDVADLRVRSS-TLKGVTVPEDR 320
           |: :| :|||||||:  :: |||::  |:    |   | :  :|| :|  |||| |||
Db     234 TVTDVGLNPTRTGLLEVMKLMGANLEWEITEENL---EPIGTVRVETSPNLKGVVVPEHL 290

Qy     321 APSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEGETSLV 380
             | |||| |::|:   |||| ||:   |||| |||||:  :    |  ||: :| :
Db     291 VPLMIDELPLVALLGVFAEGETVVRNAEELRKKESDRIRVLVENFKRLGVEIEEFKDGFK 350

Qy     381 VRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPEFMDLM 440
           : |:   |      | :|    |||:|| | : |||||  | | |   :| ||| | :|:
Db     351 IVGKQSIK------GGSVDPEGDHRMAMLFSIAGLVSEEGVDVKDHECVAVSFPNFYELL 404


RESULT  12
AROA_CAMJE
ID   AROA_CAMJE     STANDARD;      PRT;   428 AA.
AC   P52312; Q9PP36;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR CJ0895C.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC   Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=81116;
RX   MEDLINE=97128776; PubMed=8973316;
RA   Woesten M.M.S.M., Dubbink V.H.J., van der Zeijst B.A.M.;
RT   "The aroA gene of Campylobacter jejuni.";
RL   Gene 181:109-112(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X89371; CAA61554.1; -.
DR   EMBL; AL139076; CAB73153.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
FT   CONFLICT    152    152       K -> N (IN REF. 1).
FT   CONFLICT    158    158       F -> Y (IN REF. 1).
FT   CONFLICT    179    179       D -> N (IN REF. 1).
FT   CONFLICT    183    183       T -> A (IN REF. 1).
FT   CONFLICT    217    217       N -> S (IN REF. 1).
FT   CONFLICT    241    241       V -> A (IN REF. 1).
SQ   SEQUENCE   428 AA;  47288 MW;  465E2B50F39FC5AB CRC64;


  Query Match             25.8%;  Score 590;  DB 1;  Length 428;
  Best Local Similarity   34.0%;  Pred. No. 8.2e-30;
  Matches  146;  Conservative   94;  Mismatches  167;  Indels   22;  Gaps   11;

Qy      24 IPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTWIIDGVG 83
           |  |||||||  :|  |   | :    |  :| :|| : :: :||:| ::     :  :
Db      17 IAADKSISHRFAIFSLLTQEENKAQNYLLAQDTLNTLEIIKNLGAKIEQKDS--CVKIIP 74

Qy      84 NGGLLAPEAPLDFGNAATGCRLTMG-LVGVYDFDSTFI--GDASLTKRPMGRVLNPLREM 140
             :|:|    || ||: |  || :| | |:  |   |:  ||  |  ||| |:  || ::
Db      75 PKEILSPNCILDCGNSGTAMRLMIGFLAGISGF---FVLSGDKYLNNRPMRRISKPLTQI 131

Qy     141 GVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPIMT 199
           | ::   :   | |: : | |      ::  ::|||||:|::|:      : |  |  ::
Db     132 GARIYGRNEANLAPLCIEGQKLKA-FNFKSEISSAQVKTAMILSAFRADNVCTFSEISLS 190

Qy     200 RDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAALLVP 259
           |:|:| ||:   | : |  |    :    |  :  |  | | :| ||||  : ::||:::|
Db     191 RNHSENMLKAMKAPIRVSNDGLSLEINPL--KKPLKAQNIIIPNDPSSAFYFVLAAIILP 248

Qy     260 GSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVTVPED 319
           | : : |:|:|||      ||:||| :|:   :    | : ::|| || | |: | :|
Db     249 KSQIILKNILLNPTRIEAYKILQKMGAKLEMTITQ-NDFETIGEIRVESSKLNGIEV-KD 306

Qy     320 RAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEGETSL 379
             :||| | ||:| | |:| : :   :|||||||||:: :   ||| ||:  | :
Db     307 NIAWLIDEAPALAIAFALAKGKSSLINAKELRVKESDRIAVMVENLKLCGVEARELDDGF 366

Qy     380 VVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSFPEFMDL 439
           : |  : |       : : :: |||||||| ::||:    : :||:  | |||| |:::
Db     367 EIEGGCELK------SSKIKSYGDHRIAMSFAILGLLC--GIEIDDSDCIKTSFPNFIEI 418

Qy     440 MAGLGAKIE 448
           :: |||:|:
Db     419 LSNLGARID 427


RESULT  13
```

```
AROA_STAAU
ID   AROA_STAAU     STANDARD;      PRT;   430 AA.
AC   Q05615;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 8325-4;
RX   MEDLINE=93381456; PubMed=8371108;
RA   O'Connell C.M., Pattee P., Foster T.J.;
RT   "Sequence and mapping of the aroA gene of Staphylococcus aureus
RT   8325-4.";
RL   J. Gen. Microbiol. 139:1449-1460(1993).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L05004; AAA71897.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase.
SQ   SEQUENCE   430 AA;  46852 MW;  E5FC878EA1C23C20 CRC64;

 Query Match             25.2%;  Score 577;  DB 1;  Length 430;
 Best Local Similarity   33.6%;  Pred. No. 5.3e-29;
 Matches  144;  Conservative   82;  Mismatches  179;  Indels   24;  Gaps    8;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           |  | | : :|||||::||: |   || | : |   | |||   |    : :|  |:::
Db      10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGVSTIYKPLLGEDCRRTMDIFRHLGVEIKEDD 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :  ::   |   :  |   |  ||: |  ||  ||:     :|   || |: |||| |||
Db      70 EKLVVTSPGY-QVNTPHQVLYTGNSGTTTRLLAGLLSGLGNESVLSGDVSIGKRPMDRVL 128

Qy     135 NPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
            ||: |   ::  : :  |: :: |     | |:: :||||||||:| | | :   | :
Db     129 RPLKLMDANIEGIEDNYTPLIIK-PSVIKGINYQMEVASAQVKSAILFASLFSKEPTIIK 187

Qy     195 EPIMTRDHTEKMLQGF-----GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249
           |  ::|:||| | : |        |:: |  : :| |:             |||| || |
Db     188 ELDVSRNHTETMFKHFNIPIEAEGLSINTTPEAIRYIKPAD--------FHVPGDISSAA 239

Qy     250 FPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVR-S 308
           | :|||: |||||| || :| ||:|:|  :::|| :|:: | :  | |  | :|:: :
Db     240 FFIVAALITPGSDVTIHNVGINQTRSGIIDIVEKMGGNIQLFN-QTTGAEPTASIRIQYT 298

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
             |: :|:  :  |  ||| |::|:    | | : :   |||:|||::|:   |: | |

Db     299 PMLQPITIEGELVPKAIDELPVIALLCTQAVGTSTIKDAEELKVKETNRIDTTADMLNLL 358

Qy     369 GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATM 428
           | :       |::      :   ||:        |||| |   |  ::|  || :
Db     359 GFELQPTNDGLIIH---PSEFKTNATDILT----DHRIGMMLAVACVLSSEPVKIKQFDA 411

Qy     429 IATSFPEFM 437
           :  ||| |:
Db     412 VNVSFPGFL 420


RESULT  14
AROA_HELPJ
ID   AROA_HELPJ     STANDARD;      PRT;   429 AA.
AC   Q9ZKF7;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR JHP0980.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001527; AAD06557.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   429 AA;  47167 MW;  92724C4A25752741 CRC64;


 Query Match             22.8%;  Score 521.5;  DB 1;  Length 429;
 Best Local Similarity   33.6%;  Pred. No. 1.5e-25;
 Matches  145;  Conservative   73;  Mismatches  166;  Indels   47;  Gaps   13;

Qy      27 DKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG--------DTWI 78
           |||:|||: :|  ||     :   | ||| ::: :  | :||::             | |
Db      10 DKSLSHRAVIFSLLAQKPCFVRNFLMGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTTI 69

Qy      79 IDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLR 138
           :          |   |:  |: |  ||  ||:          || ||  ||| |:: ||:
Db      70 KE---------PNKILNCNNSGTSMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK 120
```

```
Qy     139 EMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEPI 197
           | ::   ||    |: : |         |  |:||||||| :|: |   ||:   |
Db     121 AFGAKILGREDNHFAPLAIVGGPLKA-CDYESPIASAQVKSAFILSALQAQGGISAYKESE 179

Qy     198 MTRDHTEKMLQGFGANLTVETDADGVRTIR-LEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
           ::|:||| ||:  |||:    : |||  |  ||    |      :  ||||  |  :|
Db     180 LSRNHTEIMLKSLGANI---QNQDGVLKISPLE--KPLESFDFTIANDPSSAFFLALACA 234

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIE-VINPRLAGGEDVADLRVRSSTLKGVT 315
           : | | : : |||:||||      |::||| || ||  :    | : |: :  : || ::
Db     235 ITPKSRLLLKNVLLNPTRIEAFEVLKKMGAHIEYVIQSK--DLEVIGDIYIEHAPLKAIS 292

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
           : :: | |:||| | |::|  ||:| :::   ::|| |||||: || :  |  |::|:|
Db     293 IDQNIA-SLIDEIPALSIAMLFAKGKSMVRNAKDLRAKESDRIKAVVSNFKALGIECEEF 351

Qy     376 ETSLVVRGRPDGKGLGNAS----------GAAVATHLDHRIAMSFLVMGLVSENPVTVDD 425
           |    :      :|||:||            : :  |||||||| |: |    |: :|:
Db     352 EDGFYI------EGLGDASQLKQHFSKIKPPIIKSFNDHRIAMSFAVLTLAL--PLEIDN 403

Qy     426 ATMIATSFPEF 436
           |||  |
Db     404 LECANISFPTF 414
```

```
RESULT  15
AROA_HELPY
ID   AROA_HELPY     STANDARD;      PRT;   429 AA.
AC   P56197;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE   enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
GN   AROA OR HP0401.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC       phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000556; AAD07470.1; -.
DR   TIGR; HP0401; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Transferase; Complete proteome.
SQ   SEQUENCE   429 AA;  47240 MW;  19545753E081FDAE CRC64;
```

Query Match 22.7%; Score 518.5; DB 1; Length 429;
Best Local Similarity 33.5%; Pred. No. 2.2e-25;
Matches 145; Conservative 74; Mismatches 171; Indels 43; Gaps 13;

```
Qy      27 DKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG--------DTWI 78
           |||:|||: :|  ||     :   | ||| ::: :  | :||::             | |
Db      10 DKSLSHRAVIFSLLAQKPCFVRNFLMGEDCLSSLEIAQNLGAKVENTAKNSFKITPPTTI 69

Qy      79 IDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPLR 138
           :          |   |:  |: |  ||  ||:          || ||  ||| |:: ||:
Db      70 KE---------PNKILNCNNSGTTMRLYSGLLSAQKGLFVLSGDNSLNARPMKRIIEPLK 120

Qy     139 EMGVQVKS-EDGDRLPVTLRGPKTPTPIT---YRVPMASAQVKSAVLLAGLNTPGITTVI 194
           | ::   ||    |: : |    :|:    |  |:|||||||| :|: |   | :|
Db     121 AFGAKILGREDNHFAPLVILG----SPLKACHYESPIASAQVKSAFILSALQAQGASTYK 176

Qy     195 EPIMTRDHTEKMLQGFGANLTVETDADGVRTIR-LEGRGKLTGQVIDVPGDPSSTAFPLV 253
           |  ::|:||| ||:  ||::    : |||  |  ||    |      :  ||||  |  :
Db     177 ESELSRNHTEIMLKSLGADI---HNQDGVLKISPLE--KPLEAFDFTIANDPSSAFFFAL 231

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIE--VINPRLAGGEDVADLRVRSSTL 311
           |  : | | : : |||:||||      |::||| ||  : :  |   | : |: |  : |
Db     232 ACAITPKSRLLLKNVLLNPTRIEAFEVLKKMGASIEYAIQSKDL---EMIGDIYVEHAPL 288

Qy     312 KGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVD 371
           | : : :: | |:||| | |::|  ||:| :::   ::|| |||||: || :  |  |::
Db     289 KAINIDQNIA-SLIDEIPALSIAMLFAKGKSMVKNAKDLRAKESDRIKAVVSNFKALGIE 347

Qy     372 CDEGETSLVVRGRPDGKGL----GNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDAT 427
           |:| |    | |  |   |           : :  |||||||| |: |    |: :|:
Db     348 CEEFEDGFYVEGLEDISPLKQRFSRIKPPLIKSFNDHRIAMSFAVLTLAL--PLEIDNLE 405

Qy     428 MIATSFPEFMDLM 440
           |||:|  |:
Db     406 CANISFPQFKHLL 418
```

Search completed: August 15, 2002, 14:06:06
Job time: 405 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:58:51 ; Search time 51.67 Seconds
(without alignments)
1523.373 Million cell updates/sec

Title: US-09-464-099A-3
Perfect score: 2288
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2104	92.0	455	16	Q92SV5	Q92sv5 rhizobium m
2	1861.5	81.4	480	2	Q9AGV2	Q9agv2 brucella ab
3	1789	78.2	452	16	Q98CC1	Q98cc1 rhizobium l
4	1133	49.5	443	16	Q9A2H2	Q9a2h2 caulobacter
5	989.5	43.2	746	16	Q9HZ69	Q9hz69 pseudomonas
6	912.5	39.9	454	16	Q9PB21	Q9pb21 xylella fas
7	884	38.6	431	16	Q9KCA6	Q9kca6 bacillus ha
8	820	35.8	430	16	Q99Z83	Q99z83 streptococc
9	815.5	35.6	428	16	Q92A85	Q92a85 listeria in
10	781.5	34.2	428	2	Q9ANY6	Q9any6 enterococcu
11	610	26.7	432	16	Q99U25	Q99u25 staphylococ
12	470.5	20.6	207	2	Q9RHZ8	Q9rhz8 pseudomonas
13	470.5	20.6	439	17	Q9HQC1	Q9hqc1 halobacteri
14	416.5	18.2	440	2	Q9L213	Q9l213 streptomyce
15	401.5	17.5	426	16	Q9KRB0	Q9krb0 vibrio chol
16	378.5	16.5	427	2	Q93ED4	Q93ed4 yersinia ru
17	351.5	15.4	428	16	Q97KM2	Q97km2 clostridium
18	332.5	14.5	438	2	Q9K4A7	Q9k4a7 streptomyce
19	324.5	14.2	447	16	Q9K9D5	Q9k9d5 bacillus ha
20	306	13.4	516	10	Q946V0	Q946v0 dicliptera
21	299.5	13.1	511	10	Q93VK6	Q93vk6 oryza sativ
22	299	13.1	444	10	O24566	O24566 zea mays (m
23	298	13.0	516	10	Q946U9	Q946u9 dicliptera
24	296	12.9	521	10	Q9FVP6	Q9fvp6 arabidopsis
25	293.5	12.8	445	8	Q95AK0	Q95ak0 eleusine in
26	292.5	12.8	445	8	Q95AK1	Q95ak1 eleusine in
27	285	12.5	433	16	Q9JTT3	Q9jtt3 neisseria m
28	280	12.2	414	17	Q980I5	Q980i5 sulfolobus
29	275	12.0	433	16	Q9JYU1	Q9jyu1 neisseria m
30	261	11.4	408	17	Q96Y91	Q96y91 sulfolobus
31	260.5	11.4	427	17	Q9YC47	Q9yc47 aeropyrum p
32	229	10.0	332	2	Q9ZEQ0	Q9zeq0 actinobacil
33	223	9.7	410	17	Q978S3	Q978s3 thermoplasm
34	216.5	9.5	391	10	O80428	O80428 oryza sativ
35	204.5	8.9	448	2	Q9L1U5	Q9l1u5 streptomyce
36	201.5	8.8	446	2	Q9S0N1	Q9s0n1 streptomyce
37	185.5	8.1	419	16	Q99Z78	Q99z78 streptococc
38	184.5	8.1	417	16	Q9K1Q9	Q9k1q9 neisseria m
39	184	8.0	418	16	Q97DD9	Q97dd9 clostridium
40	183.5	8.0	417	16	Q9JWS7	Q9jws7 neisseria m
41	182	8.0	428	16	Q9K6E5	Q9k6e5 bacillus ha
42	181	7.9	419	2	Q9S5W5	Q9s5w5 escherichia
43	179	7.8	347	10	Q9AT37	Q9at37 lolium rigi
44	179	7.8	423	16	Q927U1	Q927u1 listeria in
45	178.5	7.8	417	2	Q9EXE3	Q9exe3 mycobacteri

ALIGNMENTS

RESULT 1
Q92SV5
ID Q92SV5 PRELIMINARY; PRT; 455 AA.
AC Q92SV5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE PROTEIN (EC
DE 2.5.1.19).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoelter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL Science 293:668-672(2001).
DR EMBL; AL591783; CAC41690.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 455 AA; 47900 MW; 97659E1C7E1021B5 CRC64;

Query Match 92.0%; Score 2104; DB 16; Length 455;
Best Local Similarity 90.8%; Pred. No. 1.2e-119;
Matches 413; Conservative 18; Mismatches 24; Indels 0; Gaps 0;

Qy 1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60

```
          ||||:: ||||||||| | ||:||||||||||||||||||:||||||||||||||||||
Db      1 MSHGSNPRPATARKSSDLKGTLRIPGDKSISHRSFMFGGLAAGETRITGLLEGEDVINTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          |||||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db     61 KAMQAMGARIRKEGDTWIIDGVGNGALLAPEAPLDFGNAGTGCRLTMGLVGVYDFDSTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||||||||||:|||||||||| :|||||||||||| ||||||||||||||||||||||
Db    121 GDASLTKRPMGRVLDPLREMGVQVKSAEGDRLPVTLRGPKTPNPITYRVPMASAQVKSAV 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||||||:||:|||||||||||||||||||||||:|||||||||||||||||||||
Db    181 LLAGLNTPGITTVVEPVMTRDHTEKMLQGFGANLTVETDAEGVRTIRLEGRGKLTGQVID 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||| |:|||||:|||||||||||||||||||||||:|||:| |||||||
Db    241 VPGDPSSTAFPLVAGLIVPGSDITILNVLMNPTRTGLILTLQEMGANIEVMNKRLAGGED 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||  | |||||||||||||||||:|||||||||| ||||||||||||||||||||
Db    301 VADLRVRHSELKGVTVPEDRAPSMIDEYPVLAVAAAFAEGTTVMNGLEELRVKESDRLSA 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          ||:|||||||||||||| ||||||||  |||||   ||   | |||||||||||||||| ||:|
Db    361 VADGLKLNGVDCDEGEASLVVRGRPGGKGLGKISGGQVKTHLDHRIAMSFLVMGLASEHP 420

Qy    421 VTVDDATMIATSFPEEFMDLMAGLGAKIELSDTKAA 455
          |||||||||||||||| || ||||||| :: |||
Db    421 VTVDDATMIATSFPEEFMGLMTGLGAKIEEAENKAA 455


RESULT   2
Q9AGV2
ID   Q9AGV2      PRELIMINARY;      PRT;   480 AA.
AC   Q9AGV2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   5-ENOLPYRUVYL SHIKIMATE 3-PHOSPHATE SYNTHASE (EC 2.5.1.19).
GN   AROA.
OS   Brucella abortus.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Brucellaceae; Brucella.
OX   NCBI_TaxID=235;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2308;
RA   Gan T., Essenberg R.C.;
RT   "Characterization of the aroA gene of Brucella abortus and
RT   construction of an aroA mutant.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF326475; AAK27445.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase.
SQ   SEQUENCE   480 AA;  50634 MW;  3D55323944DA8C91 CRC64;


  Query Match           81.4%;  Score 1861.5;  DB 2;  Length 480;
  Best Local Similarity   81.7%;  Pred. No. 5.9e-105;
  Matches  365;  Conservative   35;  Mismatches   44;  Indels    3;  Gaps    1;

Qy      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||| |  :||||| |  |:| :|||||||||| ||||||||||:||||||||||||||||
Db     31 MSHSACPKPATARHSQALTGEIRIPGDKSISTRSFMFGGLASGKTRITGLLEGEDVINTG 90

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          :|||||||||||||| |||:|||||  || |||||||||  || ||||||||  ::||
Db     91 RAMQAMGARIRKEGDVWIINGVGNGCLLQPEAPLDFGNAGTGARLTMGLVGTYDMKTSFI 150

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||:|||||||||||||||:: :|||:|:|| ||:|  || ||||||||||||||||
Db    151 GDASLSKRPMGRVLNPLREMGVQVEAAEGDRMPLTLIGPRTANPIAYRVPMASAQVKSAV 210

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:|||||||||||||||||||||:|||||| |||| ||: |:|||||| ||
Db    211 LLAGLNTPGVTTVIEPVMTRDHTEKMLQGFGADLTVETDKDGVRHIRIVGQGKLTGQTID 270

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||| |||| ||:||| ||||||||||||||||||||||||:|:|||||||
Db    271 VPGDPSSTAFPLVPALLVEGSEVTIRNVLMNPTRTGLILTLQEMGADIEIIDPRLAGGED 330

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||::| |||| || :||||||||||:||: |:|||| |||:||:|||||||||:|
Db    331 VADLRVKASKLKGVVVPPERAPSMIDEYPVLAIXASFAEGETVMDGLDELRVKESDRLAA 390

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || ||||| |||||   |  |||||||||||||||||| || |
Db    391 VARGLEANGVDCTEGEMSLTVRGRPGGKGLG---GGTVATHLDHRIAMSFLVMGLASEKP 447

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKI 447
          |||||:|||||||||| :||||||||
Db    448 VTVDDSTMIATSFPEFMGMMAGLGAKI 474


RESULT   3
Q98CC1
ID   Q98CC1      PRELIMINARY;      PRT;   452 AA.
AC   Q98CC1;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   MLL5213.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003006; BAB51700.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   452 AA;  47455 MW;  2B52983E3523B938 CRC64;


  Query Match           78.2%;  Score 1789;  DB 16;  Length 452;
  Best Local Similarity   78.0%;  Pred. No. 1.3e-100;
  Matches  352;  Conservative   33;  Mismatches   62;  Indels    4;  Gaps    2;

Qy      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||| |:::|||||||  |||| |:|||||||||||||||||||||||||||||||: ||
Db      1 MSHAAAAKPATARKSQALSGTARVPGDKSISHRSFMFGGLASGETRITGLLEGEDVMRTG 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          ||:|||| | | |  |:| | ||| || || |||||| || |||||||| || ::|||
Db     61 AAMKAMGAHIEKRGAEWVIRGTGNGALLQPEGPLDFGNAGTGSRLTMGLVGTYDMETTFI 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQV-KSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSA 179
          |||||: ||||||| |||:|||||  |:  |||:|:|| |||   ||||||||||||||||
```

```
Db    121 GDASLSGRPMGRVLEPLRQMGVQVLKATPGDRMPITLHGPKHAAPITYRVPMASAQVKSA 180

Qy    180 VLLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVI 239
          ||||||||||||||||||:||||||||||:||||||:||||  ||| | :||:|:|||| |
Db    181 VLLAGLNTPGITTVIEPVMTRDHTEKMLKGFGANLSVETDERGVRHIFIEGQGRLTGQTI 240

Qy    240 DVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGE 299
          |||||||  |||||||:|||||:|| ||||||||||:|||||||  |:::|||| ||||
Db    241 AVPGDPSSAGFPLVAALIVPGSDITIENVLMNPTRTGLLLTLQEMGGQIDILNPRNAGGE 300

Qy    300 DVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLS 359
          |||||||| | |||| || :|||||||||||:|||||:|||| |:| ||||||||||||||
Db    301 DVADLRVRYSELKGVAVPPERAPSMIDEYPVLAVAASFAEGETLMQGLEELRVKESDRLS 360

Qy    360 AVANGLKLNGVDCDEGETSLVVRGRPDGKGLG---NASGAAVATHLDHRIAMSFLVMGLV 416
          |||||||||||||| ||| || ||||| |||||    |     | |||||||||||||||
Db    361 AVANGLKLNGVDCTEGEASLAVRGRPGGKGLGGHPNGLDTTVQTHLDHRIAMSFLVMGLA 420

      417 SENPVTVDDATMIATSFPEFMDLMAGLGAKI 447
          :| |||:||  |||||||||| || ||||:|
Db    421 TEKPVTIDDQAMIATSFPEFMGLMTGLGAEI 451


RESULT   4
Q9A2H2
ID   Q9A2H2      PRELIMINARY;      PRT;   443 AA.
AC   Q9A2H2;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   CC3589.
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC   Caulobacter.
OX   NCBI_TaxID=69394;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 19089 / CB15;
RX   MEDLINE=21173698; PubMed=11259647;
RA   Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA   Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA   Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA   Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA   Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA   Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT   "Complete genome sequence of Caulobacter crescentus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR   EMBL; AE006017; AAK25551.1; -.
DR   TIGR; CC3589; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   443 AA;  46075 MW;  45E8A1463E10B6EC CRC64;


  Query Match           49.5%;  Score 1133;  DB 16;  Length 443;
  Best Local Similarity  55.8%;  Pred. No. 6.4e-61;
  Matches  250;  Conservative   45;  Mismatches  137;  Indels   16;  Gaps    6;

Qy      4 GASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAM 63
          |  | |  |     | | || |||||||||| : | ||:| | : |||||:|||: | :||
Db      5 GLKSAPGGA-----LRGIVRAPGDKSISHRSMILGALATGTTTVEGLLEGDDVLATARAM 59

Qy     64 QAMGARIRKEG-DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGD 122
          || |||| :||   | |:  | ||   |   :| ||| || || ||    :   :|| ||
Db     60 QAFGARIEREGVGRWRIE--GKGGFEEPVDVIDCGNAGTGVRLIMGAAAGFAMCATFTGD 117

Qy    123 ASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLL 182
          ||  |||||||:||  ||      |  |||:||:|      : | :|||||||||||||
```

```
Db    118 QSLRGRPMGRVLDPLARMGATWLGRDKGRLPLTLKGGNL-RGLNYTLPMASAQVKSAVLL 176

Qy    183 AGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRT---IRLEGRGKLTGQVI 239
          |||:  |   ||||  ||||||:||: ||| : ||    | :|   :||    ||||  :
Db    177 AGLHAEGGVEVIEPEATRDHTERMLRAFGAEVIVEDRKAGDKTFRHVRLPEGQKLTGTHV 236

Qy    240 DVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGE 299
          |||||| ||||||||:||||:||:  |::|  |||  |||||||: : | |:| ||
Db    237 AVPGDPSSAAFPLVAALIVPGSEVTVEGVMLNELRTGLFTTLQEMGADLVISNVRVASGE 296

Qy    300 DVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLS 359
          :| |:  | | |||| || :||||||||||||||||||| | ||| |: |:||||||:|
Db    297 EVGDITARYSQLKGVVVPPERAPSMIDEYPILAVAAAFASGETVMRGVGEMRVKESDRIS 356

Qy    360 AVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSEN 419
          |||||  ||  |     :|     | |      || | || ||||||| |::|: ::
Db    357 LTANGLKACGVQVVEEPEGFIV----TGTGQPPKGGATVVTHGDHRIAMSHLILGMAAQA 412

Qy    420 PVTVDDATMIATSFPEFMDLMAGLGAKI 447
          | ||:  |||||| | ||| |||| :
Db    413 EVAVDEPGMIATSFPGFADLMRGLGATL 440


RESULT   5
Q9HZ69
ID   Q9HZ69      PRELIMINARY;      PRT;   746 AA.
AC   Q9HZ69;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   STILL FRAMESHIFT 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE
DE   PREPHENATE DEHYDROGENASE.
GN   PA3164.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004740; AAG06552.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   InterPro; IPR000205; NAD_binding.
DR   InterPro; IPR003099; PDH.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   Pfam; PF02153; PDH; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   746 AA;  79320 MW;  C2974B4BBF539E3A CRC64;


  Query Match           43.2%;  Score 989.5;  DB 16;  Length 746;
  Best Local Similarity  48.5%;  Pred. No. 6.3e-52;
  Matches  214;  Conservative   62;  Mismatches  152;  Indels   13;  Gaps    3;

Qy     12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
          |:    ||||:|:|||||||||| | | || | | : | ||||| : | :| : ||  |
Db    317 AQPGGSLSGTIRVPGDKSISHRSIMLGSLAEGTTEVEGFLEGEDALATIQAFRDMGVVIE 376

Qy     72 -KEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
           :     : |||  || ||  |:  ||: |  ||  ||:    ||||  |||||:||||
```

```
Db    377 GPQNGRVTVHGVGLHGLKAPPGPIYLGNSGTSMRLLSGLLAAQPFDSTLTGDASLSKRPM 436

Qy    131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
          ||  |||||| :::    | |:|:|| :  | : | :|||||||||| :|||||   |
Db    437 NRVAKPLREMGAVIETGPEGRPPMTIRGGQRLTGMHYDMPMASAQVKSCLLLAGLYAAGE 496

Qy    191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
          |:| ||  |||||||:||:|||  : ||       | ::|   ||:   |:|| | || ||
Db    497 TSVTEPAPTRDHTERMLRGFGYPVVVEGS-----TAKVESGHKLSATHIEVPADISSAAF 551

Qy    251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
          |||| :  ||:: : :| :|||| |:|  |: || |: : | |  ||| |||:|||||:
Db    552 FLVAASIAEGSELVLQHVGINPTRVGVIEILRLMGGDLSLENQREVGGEPVADIRVRSAR 611

Qy    311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
          |||: :|||  |  |||:|:| |||| ||| ||: | ||||||||||:  :|:|||  ||
Db    612 LKGIDIPEDLVPLAIDEFPVLFVAAACAEGRTVLRGAEELRVKESDRIQVMADGLKALGV 671

Qy    371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
          :    :|: |       |  | | | ||||||||| |  | :  |: : |   :|
Db    672 KAEPTPDGIVIEG-------GAFGGGEVWAHGDHRIAMSFSVASLRASGPIRIHDCANVA 724

Qy    431 TSFPEFMDLMAGLGAKIELSD 451
          |||| |: | |  | :: : :
Db    725 TSFPNFLALCAQTGIRVAVEN 745


RESULT   6
Q9PB21
ID   Q9PB21      PRELIMINARY;      PRT;   454 AA.
AC   Q9PB21;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   XF2324.
OS   Xylella fastidiosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC   Xylella.
OX   NCBI_TaxID=2371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=9A5C;
RX   MEDLINE=20365717; PubMed=10910347;
RA   Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA   Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA   Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA   Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA   Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA   Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA   Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA   Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA   Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA   Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA   Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA   Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA   Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA   Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA   Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA   Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA   Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA   de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA   Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA   Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA   de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA   da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA   da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA   de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,
RA   Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA   Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT   "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL   Nature 406:151-159(2000).
DR   EMBL; AE004043; AAF85123.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Complete proteome.
SQ   SEQUENCE   454 AA;  48266 MW;  45CCF074E6C0BA57 CRC64;


  Query Match             39.9%;  Score 912.5;  DB 16;  Length 454;
  Best Local Similarity   46.9%;  Pred. No. 1.5e-47;
  Matches  206;  Conservative   60;  Mismatches  158;  Indels   15;  Gaps    4;

Qy     12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
          | : : | | : ||||||||||: ||  || | :|| | || ||  :| : :  :| ||
Db     23 AHQGTPLHGVLSIPGDKSISHRAVMFAALADGTSRIDGFLEAEDTCSTAEILARLGVRIE 82

Qy     72 KEGDTW-IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPM 130
             |  |: ||| || |  ||| ||| || ||  ||:    |||  :|||||:|||| 
Db     83 TPLSTQRIVHGVGVDGLQASHIPLDCGNAGTGMRLLAGLLVAQPFDSVLVGDASLSKRPM 142

Qy    131 GRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
          || :|| :|| :: : |    |: : | :    | :  |:||||:||||||||| 
Db    143 RRVTDPLSQMGARIDTSDDGTPPLRIYGGQLLHGIDFISPVASAQIKSAVLLAGLYARNE 202

Qy    191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
          | | ||  |||:||:||  || :: | |        || |  :|    | :| | || ||
Db    203 TVVREPHPTRDYTERMLTAFGVDIDVSTGC-----ARLRGGQRLCATDITIPADFSSAAF 257

Qy    251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSST 310
          |||| ::|||||:|:  | :|| | ||:  |: |||:|   |    ||| | ||||| :
Db    258 YLVAASVIPGSDITLRAVGLNPRRIGLLTVLRLMGANIVESNRHEQGGEPVVDLRVRYAP 317

Qy    311 LKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGV 370
          |:|  ||||    ||||:| | |||| ||| ||::|  ||||||||||:|:  ||:: ||
Db    318 LQGTRVPEDLVADMIDEFPALFVAAAAAEGQTVVSGAAELRVKESDRLAAMVTGLRVLGV 377

Qy    371 DCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIA 430
            ||      : | | | |  |: |       ||||||:| : | :| : | ::|   :|
Db    378 QVDETADGATIHGGPIGHGTINSHG-------DHRIAMAFSIAGQLSVSTVRIEDVANVA 430

Qy    431 TSFPEFMDL--MAGLGAKI 447
          ||||::  |   || | ::
Db    431 TSFPDYETLARSAGFGLEV 449


RESULT   7
Q9KCA6
ID   Q9KCA6      PRELIMINARY;      PRT;   431 AA.
AC   Q9KCA6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE   ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)
DE   (EPSPS).
GN   AROE OR BH1667.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
CC   -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =
CC       ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
```

```
CC   -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC       AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
DR   EMBL; AP001512; BAB05386.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ   SEQUENCE   431 AA;  45485 MW;  12F4FFBE7BA0743D CRC64;

  Query Match             38.6%;  Score 884;  DB 16;  Length 431;
  Best Local Similarity   44.2%;  Pred. No. 7.2e-46;
  Matches  188;  Conservative   72;  Mismatches  153;  Indels   12;  Gaps    5;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
           :  || ||:::|||||||||: ||| || | | : | | | | ::|    | :|  | :
Db      10 AKGLKGTIKVPGDKSISHRAVMFGALAKGTTTVEGFLPGADCLSTISCFQKLGVSIEQAE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
           :   : | |  ||  |   || ||: |  || :|::    | |  ||| |: |||| ||
Db      70 ERVTVKGKGWDGLREPSDILDVGNSGTTTRLILGILSTLPFHSVIIGDESIGKRPMKRVT 129

Qy     135 NPLREMGVQVKSED-GDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTV 193
           ||: || |:   | |:  |:::|| :    | :  |:||||:|||:|||||   | |:|
Db     130 EPLKSMGAQIDGRDHGNLTPLSIRGGQL-KGIDFHSPVASAQMKSAILLAGLRAEGKTSV 188

Qy     194 IEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLV 253
           ||  ||||||:||: || |:     : ||: |: :||   |||| : |||| || || ||
Db     189 TEPAKTRDHTERMLEAFGVNI----EKDGL-TVSIEGGQMLTGQHVVVPGDISSAAFFLV 243

Qy     254 AALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKG 313
           |  :|| | :|: || :|||| |::  |::||| : : | |: ||| |||| : :| |:|
Db     244 AGAMVPHSRITLTNVGINPTRAGILEVLKQMGATLAMENERVQGGEPVADLTIETSVLQG 303

Qy     314 VTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCD 373
           | :  |  | :||| ||:|| |  | | ||:   |||:|||::|:  | : |   |
Db     304 VEIGGDIIPRLIDEIPIIAVLATQASGRTVIKDAEELKVKETNRIDTVVSELTKLGASIH 363

Qy     374 EGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATMIATSF 433
           :. ::: |    ||     |  |::| |||| |:  :  |::| ||||:    || |:
Db     364 ATDDGMIIEGPTPLKG-----GVTVSSHGDHRIGMAMAIAALLAEKPVTVEGTEAIAVSY 418

Qy     434 PEFMD 438
           | | |
Db     419 PSFFD 423


RESULT   8
Q99Z83
ID   Q99Z83      PRELIMINARY;      PRT;   430 AA.
AC   Q99Z83;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   PUTATIVE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE
DE   (EC 2.5.1.19).
GN   AROA OR SPY1352.
OS   Streptococcus pyogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1314;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX   MEDLINE=21192684; PubMed=11296296;
RA   Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA   Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA   Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J.,
RA   Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT   "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR   EMBL; AE006573; AAK34180.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   430 AA;  46692 MW;  3273C6B39020FB61 CRC64;

  Query Match             35.8%;  Score 820;  DB 16;  Length 430;
  Best Local Similarity   43.8%;  Pred. No. 5.4e-42;
  Matches  189;  Conservative   76;  Mismatches  141;  Indels   26;  Gaps    9;

Qy      13 RKSSG-LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
           | ::| | ||:::|||||||||: : | :| |||| |||:||||::| :| : :| ||
Db       7 RTNAGPLQGTIQVPGDKSISHRAVILGAVAKGETRVKGLLKGEDVLSTIQAFRNLGVRIE 66

Qy      72 KEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMG 131
           :: |  :|:| |  || ||   |: ||: |  ||  ||:    |   ||| ||:||||
Db      67 EKDDQLVIEGQGFQGLNAPCQTLNMGNSGTSMRLIAGLLAGQPFSVKMIGDESLSKRPMD 126

Qy     132 RVLNPLREMGVQVKSE-DGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGI 190
           |:: ||::|||::  | |    |: |:| :   |||| :|::|||||||:||| |   |
Db     127 RIVYPLKQMGVEISGETDRQFPPLQLQGNRNLQPITYTLPISSAQVKSAILLAALQAKGT 186

Qy     191 TTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAF 250
           | |:|  :||:|||:|:| ||  | |    || | | | |  :|| | | |||| || ||
Db     187 TQVVEKEITRNHTEEMIQQFGGRLIV----DGKR-ITLVGPQQLTAQEITVPGDISSAAF 241

Qy     251 PLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADI--EVINPRLAGGEDVADLRVRS 308
           ||| |::|||:: : || :|||||::  :::||| |  | :| :    | |  :||
Db     242 WLVAGLIIPGSELLLKNVGVNPTRTGILEVVEKMGAQIVYEDMNKK----EQVTSIRVVY 297

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
           | :||  :     | :||| ||:|: |  |:| | :   :||||||:||:  | : |
Db     298 SNMKGTIISGGLIPRLIDELPIIALLATQAQGTTCIKDAQELRVKETDRIQVVTDILNSM 357

Qy     369 GVDCDEGETSLVVRGRPDG---KGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVD 424
           | :         ::   ||   ||     ||  :|: |||| |   :   || :  | :|
Db     358 GAN---------IKATADGMIIKGPTVLYGANTSTYGDHRIGMMTAIAALLVKQGQVHLD 408

Qy     425 DATMIATSFPEF 436
           | ||:| |
Db     409 KEEAIMTSYPTF 420


RESULT   9
Q92A85
ID   Q92A85      PRELIMINARY;      PRT;   428 AA.
AC   Q92A85;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   AROE PROTEIN.
GN   AROE OR LIN2037.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
```

```
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596170; CAC97267.1; -.
DR   ListiList; LIN02037; -.
KW   Complete proteome.
SQ   SEQUENCE   428 AA;  45994 MW;  157B48C091A68FEB CRC64;

  Query Match             35.6%;  Score 815.5;  DB 16;  Length 428;
  Best Local Similarity   40.8%;  Pred. No. 1e-41;
  Matches  173;  Conservative   83;  Mismatches  155;  Indels   13;  Gaps    5;

Qy      17 GLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDT 76
           || | : :|||||:|||| ||| :| |:| |   |  :| : | || :|:| :| :  :
Db       9 GLVGEITVPGDKSMSHRSIMFGAIAEGKTVIRHFLRADDCLGTIKAFKALGVKIEETDEE 68

Qy      77 WIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNP 136
             |: | |: ||   | ||| ||: |  || ||::   |||:  :|| |: |||| ||: |
Db      69 IIVHGTGSDGLKQAEGPLDIGNSGTTIRLMMGILAGRDFDTVILGDESIAKRPMNRVMLP 128

Qy     137 LREMGVQVKSEDGDRL-PVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIE 195
           |:||| ::  :||    |::: | ::   : | :|:|||||||:: | |   | | : |
Db     129 LQEMGAKMHGKDGSEFAPISIIGNQSLKRMEYHMPVASAQVKSAIIFAALQAEGETIIHE 188

Qy     196 PIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAA 255
              |||||| |:: ||     | : ||: |||::|  |  || : |||| || || :||
Db     189 KEKTRDHTEHMIRQFGG----EIEMDGL-TIRVKGGQKFIGQEMTVPGDVSSAAFFIVAG 243

Qy     256 LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVRSSTLKGVT 315
           |: |||:: : :| :|||||:   :::|| : | :   : |:    : |:|| |||
Db     244 LITPGSEIELTHVGLNPTRTGIFDVVEQMGGSLVVKDSSRSTGKLAGTVVVKSSELKGTE 303

Qy     316 VPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDCDEG 375
           :  |  | :||| |::|: |  ||| |::    ||:|||::|: |||| |   | |
Db     304 IGGDIIPRLIDEIPVIALLATQAEGTTIIKDAAELKVKETNRIDAVANELNKMGADITPT 363

Qy     376 ETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIATSFP 434
           |  |::||:           | | :: |||| |   :   || :  | :| |  :: |:|
Db     364 EDGLIIRGKTP------LHAANVTSYGDHRIGMMLQIAALLVEDGDVELDRAEAVSVSYP 417

Qy     435 EFMD 438
             | :
Db     418 TFFE 421


RESULT  10
Q9ANY6
ID   Q9ANY6      PRELIMINARY;      PRT;   428 AA.
AC   Q9ANY6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19).
GN   AROA.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC   Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OG1RF;
RA   Huycke M.M., Shepard L., Joyce W., Wise P., Moore D.R., Gilmore M.S.;
RT   "Forme fruste respiration by Enterococcus faecalis produces
RT   extracellular superoxide.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF318277; AAG53678.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase.
SQ   SEQUENCE   428 AA;  45715 MW;  9E0F4FE4A893CA95 CRC64;

  Query Match             34.2%;  Score 781.5;  DB 2;  Length 428;
  Best Local Similarity   40.5%;  Pred. No. 1.1e-39;
  Matches  173;  Conservative   79;  Mismatches  152;  Indels   23;  Gaps    7;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGDTW 77
           | ||: :| |||||||| ||| ::||:| ||  | ||| ::|  | :::|  |  :| |
Db      12 LQGTLMVPSDKSISHRSIMFGAISSGKTTITNFLRGEDCLSTLAAFRSLGVNIEDDGTTI 71

Qy      78 IIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLNPL 137
             ::| |  ||   :  :| ||: |  || :|::    |::   |||| |||| ||: ||
Db      72 TVEGRGFAGLKKAKNTIDVGNSGTTIRLMLGILAGCPFETRLAGDASIAKRPMNRVMLPL 131

Qy     138 REMGVQVKS-EDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVIEP 196
             :|| : :  :  :  |:::|| :   || | :|:||||||||:| | |   | : |:|
Db     132 NQMGAECQGVQQTEFPPISIRGTQNLQPIDYTMPVASAQVKSAILFAALQAEGTSVVVEK 191

Qy     197 IMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTAFPLVAAL 256
             ||||||:|:: ||  | |    || : | | |  :|||| : |||| || || ||| |
Db     192 EKTRDHTEEMIRQFGGTLEV----DG-KKIMLTGPQQLTGQNVVVPGDISSAAFFLVAGL 246

Qy     257 LVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVA----DLRVRSSTLK 312
           :|| |:: : || :| ||||::  :: ||  : ::|      || |    || |::| |
Db     247 VVPDSEILLKNVGLNQTRTGILDVIKNMGGSVTILN------EDEANHSGDLLVKTSQLT 300

Qy     313 GVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLNGVDC 372
           :     | :||| ||:|: |  | | |::   |||:|||::|: |||  | : | |
Db     301 ATEIGGAIIPRLIDELPIIALLATQATGTTIIRDAEELKVKETNRIDAVAKELTILGADI 360

Qy     373 DEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMG-LVSENPVTVDDATMIAT 431
           :  |:: |       :  |  | :: |||| |   :   || |  | :| |  ::
Db     361 TPTDDGLIIHGPT------SLHGGRVTSYGDHRIGMMLQIAALLVKEGTVELDKAEAVSV 414

Qy     432 SFPEFMD 438
           |:| | |
Db     415 SYPAFFD 421


RESULT  11
Q99U25
ID   Q99U25      PRELIMINARY;      PRT;   432 AA.
AC   Q99U25;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   AROA OR SA1297 OR SAV1464.
OS   Staphylococcus aureus (strain N315), and
OS   Staphylococcus aureus (strain Mu50).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
OX   NCBI_TaxID=158879, 158878;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX   MEDLINE=21311952; PubMed=11418146;
RA   Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA   Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA   Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA   Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA   Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA   Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA   Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT   "Whole genome sequencing of meticillin-resistant Staphylococcus
RT   aureus.";
RL   Lancet 357:1225-1240(2001).
DR   EMBL; AP003134; BAB42557.1; -.
DR   EMBL; AP003362; BAB57626.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
```

```
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   432 AA;  47068 MW;  A42102057AD15C72 CRC64;


  Query Match             26.7%;  Score 610;  DB 16;  Length 432;
  Best Local Similarity   34.0%;  Pred. No. 2.8e-29;
  Matches  146;  Conservative   80;  Mismatches  181;  Indels   22;  Gaps    6;

Qy      15 SSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEG 74
Db      10 SGPLKGEIEVPGDKSMTHRAIMLASLAEGTSNIYKPLLGEDCRRTMDIFRLLGVDIKEDE 69

Qy      75 DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVL 134
Db      70 DKLVVNSPGYKAFKTPHQVLYTGNSGTTTRLLAGLLSGLGIESVLSGDVSIGKRPMDRVL 129

Qy     135 NPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTPGITTVI 194
Db     130 RPLKLMDANIEGIEDNYTPLIIK-PSVIKGINYQMEVASAQVKSAILFASLFSNDTTVIK 188

Qy     195 EPIMTRDHTEKMLQGF-----GANLTVETDADGVRTIRLEGRGKLTGQVIDVPGDPSSTA 249
Db     189 ELDVSRNHTETMFRHFNIPIEAERLSITTTPDAIQHIKPAD--------FHVPGDISSAA 240

Qy     250 FPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGEDVADLRVR-S 308
Db     241 FFIVAALITPESDVTIHNVGINPTRSGIIDIVEKMGGNIQLFN-QTTGAEPTASIRIQYT 299

Qy     309 STLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAVANGLKLN 368
Db     300 PMLQPITIEGELVPKAIDELPVIALLCTQAVGTSTIKDAEELKVKETNRIDTTADMLNLL 359

Qy     369 GVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPVTVDDATM 428
Db     360 GFELQPTNDGLIIHPSE------FKTNATVDSLTDHRIGMMLAVASLLSSEPVKIKQFDA 413

Qy     429 IATSFPEFM 437
Db     414 VNVSFPGFL 422


RESULT  12
Q9RHZ8
ID   Q9RHZ8      PRELIMINARY;      PRT;   207 AA.
AC   Q9RHZ8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   5-ENOLPYRUVYLSHIKMATE 3-P SYNTHASE (FRAGMENT).
GN   AROF.
OS   Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=316;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JM300;
RX   MEDLINE=92013931; PubMed=1919506;
RA   Fischer R.S., Zhao G., Jensen R.A.;
RT   "Cloning, sequencing, and expression of the P-protein gene (pheA) of
RT   Pseudomonas stutzeri in Escherichia coli: implications for
RT   evolutionary relationships in phenylalanine biosynthesis.";
RL   J. Gen. Microbiol. 137:1293-1301(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JM300;
RX   MEDLINE=99298294; PubMed=10368439;
RA   Xie G., Bonner C.A., Jensen R.A.;
RT   "A probable mixed-function supraoperon in Pseudomonas exhibits gene
RT   organization features of both intergenomic conservation and gene
RT   shuffling.";
RL   J. Mol. Evol. 49:108-121(1999).
DR   EMBL; AF038578; AAD47363.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
FT   NON_TER     207    207
SQ   SEQUENCE   207 AA;  21780 MW;  7231191C72A21D6B CRC64;


  Query Match             20.6%;  Score 470.5;  DB 2;  Length 207;
  Best Local Similarity   51.5%;  Pred. No. 2.8e-21;
  Matches  103;  Conservative   22;  Mismatches   70;  Indels    5;  Gaps    2;

Qy      12 ARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIR 71
Db      10 ANPGGSLTGQLRVPGDKSISHRSIMLGSLAEGTTEXEGFLEGEDALATXQAFRXMGVVI- 68

Qy      72 KEG---DTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKR 128
Db      69 -EGPHQGRVTVHGVGLHGLQAPPGPIYLGNSGTSMRLLAGLLAAQPFDTTLSGDASLTKR 127

Qy     129 PMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAVLLAGLNTP 188
Db     128 PMNRVAKPLREMGAVIETAAEGRPPLTIRGGKKLSGMHYDMPMASAQVKSCLLLAGLYAA 187

Qy     189 GITTVIEPIMTRDHTEKMLQ 208
Db     188 GKTSVTEPAPTRDHTERMQQ 207


RESULT  13
Q9HQC1
ID   Q9HQC1      PRELIMINARY;      PRT;   439 AA.
AC   Q9HQC1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN   PSC OR VNG1232G.
OS   Halobacterium sp. (strain NRC-1).
OC   Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC   Halobacterium.
OX   NCBI_TaxID=64091;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR   EMBL; AE005049; AAG19594.1; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; UNKNOWN_1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; UNKNOWN_1.
KW   Transferase; Complete proteome.
SQ   SEQUENCE   439 AA;  44333 MW;  48CAD75A1F0ECB89 CRC64;


  Query Match             20.6%;  Score 470.5;  DB 17;  Length 439;
  Best Local Similarity   31.3%;  Pred. No. 7.8e-21;
  Matches  139;  Conservative   76;  Mismatches  190;  Indels   39;  Gaps   13;
```

```
Qy      16 SGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD 75
Db      17 SRVRGRARAPPSKSYTHRALLAAGYADGETVVRDPLVSADTRATARAVELLGGAAARENG 76

Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
Db      77 DWVVTGFGSRPAI-PDAVIDCANSGTTMRLVTAAAALADGTTVLTGDESLRARPHGPLLD 135

Qy     136 PLREMGVQVKSEDGD-RLPVTLRGPKTPTPITYRVP-MASAQVKSAVLLAGLNT-----P 188
Db     136 ALSGLGGTARSTRGNGQAPLVVDGPVSGGSVA--LPGDVSSQFVTALLMAGAVTETGIET 193

Qy     189 GITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRG----KLTGQVIDVPGD 244
Db     194 DLTTELKSAPYVDITLDVLDAFGVGAS-ET-AAGYRV-----RGGQAYAPSGAEYAVPGD 246

Qy     245 PSSTAFPLVAALLVPGSDVTILNVLMNPTRTG---LILTLQEMGADIEVINPRLAGGEDV 301
Db     247 FSSASYLLAAGALAAADGAAVVVEGMHPSAQGDAAIVDVLERMGADID-------WDTES 299

Qy     302 ADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSAV 361
Db     300 GVITVQRSELSGVEVGVADTPDLL---PTIAVLGAAADGTTRITDAEHVRYKETDRVAAM 356

Qy     362 ANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENPV 421
Db     357 AESLSKLGASVEERPDELVVRG-----GDTELSGASVDGRGDHRLVMALAVAGLVADGET 411

Qy     422 TVDDATMIATSFPEFMDLMAGLGA 445
Db     412 TIAGSEHVDVSFPDFFEVLAGLGA 435
```

RESULT 14
Q9L213
ID Q9L213 PRELIMINARY; PRT; 440 AA.
AC Q9L213;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE.
GN AROA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL138598; CAB71266.1; -.
DR InterPro; IPR001986; EPSP_syntase.
DR InterPro; IPR000408; RCC1.
DR Pfam; PF00275; EPSP_syntase; 1.
DR ProDom; PD001867; EPSP_syntase; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW Transferase.

SQ SEQUENCE 440 AA; 46529 MW; 77B720F81398EB0D CRC64;

Query Match 18.2%; Score 416.5; DB 2; Length 440;
Best Local Similarity 28.0%; Pred. No. 1.5e-17;
Matches 128; Conservative 78; Mismatches 198; Indels 53; Gaps 13;

```
Qy      11 TARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMG--- 67
Db      15 TIRTLTGFDETVRVLGSKSYTNRYLAIASLSGQETVIDNALLSDDTVYFSRAIETFGHVT 74

Qy      68 -------ARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
Db      75 CDIDHATARIR-------VTPTGR-PMRAPSEDIFVGGAGTPLRFLISMAGHADGTTIIT 126

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGD-RLPVTLRGPKTPTPITYRVPMASAQVKSA 179
Db     127 GNARMQERPMGDLLKALPALGVDATAVRGNGSPPVRVVGGSFKGGATSISGAVSSQFTSS 186

Qy     180 VLLAGLNTPGIT--TVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQ 237
Db     187 LIINALRAQTDTEITISDDLVSKPYVEMTLAGL-AEMGVSVDRDGYRRFTVPSGQQARGG 245

Qy     238 VIDVPGDPSSTAFPLVAA------LLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVI 291
Db     246 QVTVEPDASGMSYFLAAAAILQSRVVIPG-----IGAGSHQGDVHLVQALERMGCRTEV- 299

Qy     292 NPRLAGGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELR 351
Db     300 ------GDD--SITVTGGPLRGIDIDMEAMPDVV---PSLAIVAAYAEGTTRITNIASLR 348

Qy     352 VKESDRLSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFL 411
Db     349 VKECDRIAAVTTELRKMGIDVEEHSDAMYITG-------GTPHGAVIDTYDDHRIAMTFA 401

Qy     412 VMGLVSENPVTVDDATMIATSFPEFMDLMAGLGAKIE 448
Db     402 IGGLRTEG-VVIKDPGCVAKSFPAFWQTLDTLHPDLE 437
```

RESULT 15
Q9KRB0
ID Q9KRB0 PRELIMINARY; PRT; 426 AA.
AC Q9KRB0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-
DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE)
DE (EPSPS).
GN VC1732.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =
CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.
CC -!- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

```
DR   EMBL; AE004251; AAF94882.1; -.
DR   TIGR; VC1732; -.
DR   InterPro; IPR001986; EPSP_syntase.
DR   Pfam; PF00275; EPSP_syntase; 1.
DR   ProDom; PD001867; EPSP_syntase; 1.
DR   PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR   PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW   Aromatic amino acid biosynthesis; Complete proteome; Transferase.
SQ   SEQUENCE   426 AA;  46101 MW;  38852D6483BFE1C3 CRC64;


  Query Match             17.5%;  Score 401.5;  DB 16;  Length 426;
  Best Local Similarity   28.0%;  Pred. No. 1.1e-16;
  Matches  125;  Conservative   87;  Mismatches  180;  Indels   55;  Gaps   15;

Qy      18 LSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTGKAMQAMGARIRKEGD-- 75
           :|| | :|| ||:|:|: :   |||| ||:| ||: :|: :   |:  :|   |   |
Db      12 ISGEVNLPGSKSVSNRALLLAALASGTTRLTNLLDSDDIRHMLNALTKLGVNYRLSADKT 71

Qy      76 TWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFIGDASLTKRPMGRVLN 135
           |  ::|:|          |  ||| |  |     : :   |    |:  : :||:| :::
Db      72 TCEVEGLGQAFHTTQPLELFLGNAGTAMRPLAAALCLGQGDYVLTGEPRMKERPIGHLVD 131

Qy     136 PLREMGVQVKSEDGDRLP------------VTLRGPKTP---TPITYRVPMASAQVKSA 179
           ||: | |::  : :  |              ||: |  :    |      |:|  :|
Db     132 ALRQAGAQIEYLEQENFPPLRIQGTGLQAGTVTIDGSISSQFLTAFLMSAPLAQGKVTIK 191

Qy     180 VLLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLT-GQV 238
           ::   :: | |          | |  ::: ||  :    : |    :   |:  :: ||
Db     192 IVGELVSKPYI----------DITLHIMEQFGVQV---INHDYQEFVIPAGQSYVSPGQF 238

Qy     239 IDVPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLIL--TLQEMGADIEVINPRLA 296
           : | || || ::  :||  : | :| :  :  |  :  :     |::||| ||
Db     239 L-VEGDASSASY-FLAAAAIKGGEVKVTGIGKNSIQGDIQFADALEKMGAQIE------- 289

Qy     297 GGEDVADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESD 356
           |:|   :  |   |  | :  :  |   |    :|  | ||:| | :  :   ||||:|
Db     290 WGDDY--VIARRGELNAVDLDFNHIP---DAAMTIATTALFAKGTTAIRNVYNWRVKETD 344

Qy     357 RLSAVANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLV 416
           ||:|:|  |:  |   :|||  :|:   |  | :     ||: |: |||:|| | :: |
Db     345 RLAAMATELRKVGATVEEGEDFIVI--TPPTKLI----HAAIDTYDDHRMAMCFSLVAL- 397

Qy     417 SENPVTVDDATMIATSFPEFMDLMAGL 443
           |: |||::|    : :||:: |  | |
Db     398 SDTPVTINDPKCTSKTFPDYFDKFAQL 424


Search completed: August 15, 2002, 14:05:43
Job time: 412 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 15, 2002, 13:57:41 ; Search time 56.25 Seconds
(without alignments)
898.464 Million cell updates/sec

Title: US-09-464-099A-3
Perfect score: 2288
Sequence: 1 MSHGASSRPATARKSSGLSG.....FMDLMAGLGAKIELSDTKAA 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2288	100.0	455	13	AAR22300	Class II EPSPS enz
2	2288	100.0	455	18	AAW34683	Class II EPSP synt
3	2288	100.0	455	18	AAW24474	Class II EPSPS for
4	2288	100.0	455	19	AAW71609	Agrobacterium sp.
5	2288	100.0	455	22	AAE05053	Agrobacterium sp.
6	2282	99.7	455	13	AAR26449	CP4-EPSPS. Synthe
7	2282	99.7	455	19	AAW39426	CP4-EPSPS protein.
8	2282	99.7	527	22	AAM52214	EPSPS SEQ ID NO 3.
9	1906.5	83.3	449	13	AAR22301	Class II EPSPS enz
10	1906.5	83.3	449	13	AAR22302	Class II EPSPS enz
11	1906.5	83.3	449	18	AAW34684	Class II EPSP synt
12	1906.5	83.3	449	18	AAW34685	Class II EPSP synt
13	1906.5	83.3	449	18	AAW24479	Class II EPSPS use
14	1906.5	83.3	449	18	AAW24480	Class II EPSPS use
15	1906.5	83.3	449	19	AAW71611	Pseudomonas sp. st
16	1906.5	83.3	449	19	AAW71610	Achromobacter sp.
17	1906.5	83.3	449	22	AAE05054	Achromobacter sp.
18	1906.5	83.3	449	22	AAE05055	Pseudomonas sp. st
19	867.5	37.9	447	18	AAW34688	Class II EPSP synt
20	867.5	37.9	447	18	AAW24487	Class II EPSPS for
21	867.5	37.9	447	19	AAW71619	Synechocystis sp.
22	867.5	37.9	447	22	AAE05072	Synechocystis sp.
23	806	35.2	443	18	AAW34689	Class II EPSP synt
24	806	35.2	443	18	AAW24488	Class II EPSPS for
25	806	35.2	443	19	AAW71620	Dichelobacter nodo
26	806	35.2	443	22	AAE05073	Dichelobacter nodo
27	803.5	35.1	427	22	AAB48177	S. pneumoniae aroA
28	803.5	35.1	427	22	AAB37094	Streptococcus pneu
29	803.5	35.1	427	22	AAB48500	Streptococcus pneu
30	798.5	34.9	427	20	AAW97388	5-enolpyruvylshiki
31	798.5	34.9	427	21	AAY56504	Streptococcus pneu
32	798.5	34.9	427	22	AAB31158	A 5-enolpyruvoylsh
33	785	34.3	428	18	AAW34686	Class II EPSP synt
34	785	34.3	428	18	AAW24481	Class II EPSPS use
35	785	34.3	428	19	AAW71617	Bacillus subtilis
36	785	34.3	428	22	AAE05070	Bacillus subtilis
37	704.5	30.8	415	20	AAW97389	Protein encoded by
38	704.5	30.8	415	21	AAY56505	Streptococcus pneu
39	704.5	30.8	415	22	AAB31159	A 5-enolpyruvoylsh
40	704.5	30.8	415	22	AAB48178	S. pneumoniae aroA
41	704.5	30.8	415	22	AAB37095	Streptococcus pneu
42	704.5	30.8	415	22	AAB48501	Protein encoded by
43	577	25.2	430	18	AAW34687	Class II EPSP synt
44	577	25.2	430	18	AAW24482	Class II EPSPS use
45	577	25.2	430	19	AAW71618	Staphylococcus aur

ALIGNMENTS

RESULT 1
AAR22300
ID AAR22300 standard; Protein; 455 AA.
XX
AC AAR22300;
XX
DT 03-AUG-1992 (first entry)
XX
DE Class II EPSPS enzyme.
XX
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX
OS Agrobacterium sp. strain CP4.
XX
PN WO9204449-A.
XX
PD 19-MAR-1992.
XX
PF 28-AUG-1991; 91WO-US06148.
XX
PR 31-AUG-1990; 90US-0576537.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 3; 148pp; English.

```
XX
CC   The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC   phosphate synthase enzyme (EPSPS) of Agrobacterium sp. strain CP4
CC   It is used to create glyphosate resistant plants or seeds which
CC   can be planted in a field of crops to selectively control weeds.
CC   The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC   tobacco and alfalfa. This provides a cost effective, environmentally
CC   compatible weed control device. See also AAR22301 and AAR22302.
XX
SQ   Sequence   455 AA;

  Query Match            100.0%;  Score 2288;  DB 13;  Length 455;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-185;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


ESULT   2
AW34683
ID   AAW34683 standard; Protein; 455 AA.
XX
AC   AAW34683;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Agrobacterium sp. strain CP4.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Agrobacterium sp. strain CP4.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          271..274
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93788.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Claim 7; Fig 3; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence is from
CC   Agrobacterium sp. strain CP4. The EPSPS enzymes are part of the shikimic
CC   acid pathway, which leads to the biosynthesis of aromatic compounds.
CC   EPSPS converts phosphoenolpyruvic acid (PEP) and 3-phosphoshikimic acid
CC   to 5-enolpyruvyl-3-phosphoshikimic acid, and is inhibited by the
CC   herbicide glyphosate. It would be useful to produce transgenic crops
CC   containing glyphosate resistance genes so that glyphosate-containing
CC   herbicides can be applied to selectively kill weeds. The novel EPSPS
CC   enzymes exhibit a low Km for PEP and a high Ki for glyphosate, such that
CC   when introduced into a plant, the plant is made glyphosate-tolerant, and
CC   EPSPS enzyme activity is not affected. These class II EPSPS enzymes are
CC   fused to a chloroplast transit peptide to target the protein into the
CC   chloroplast, which is the site for the shikimic acid pathway. In
CC   addition, the EPSPS gene is cloned into a plant under the control of a
CC   promoter such as figwort mosaic virus promoter or the cauliflower mosaic
CC   virus promoter, so that expression is enhanced.
XX
SQ   Sequence   455 AA;

  Query Match            100.0%;  Score 2288;  DB 18;  Length 455;
  Best Local Similarity  100.0%;  Pred. No. 6.2e-185;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240
```

```
Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

RESULT 3
AAW24474
ID AAW24474 standard; Protein; 455 AA.
XX
AC AAW24474;
XX
DT 02-OCT-1997 (first entry)
XX
DE Class II EPSPS for glyphosate resistant plant production.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW tobacco.
XX
OS Agrobacterium sp. strain CP4.
XX
PN US5633435-A.
XX
PD 27-MAY-1997.
XX
PF 31-AUG-1990; 90US-0576537.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1997-297418/27.
DR N-PSDB; AAT77313.
XX
PT New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT used for transforming plants to produce plants which are tolerant to
PT glyphosate herbicide
XX
PS Claim 1; Column 57-60; 154pp; English.
XX
CC AAW24474 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC the production of herbicide resistant (glyphosate-tolerant) plants
CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ Sequence 455 AA;

Query Match 100.0%; Score 2288; DB 18; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.2e-185;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

RESULT 4
AAW71609
ID AAW71609 standard; Protein; 455 AA.
XX
AC AAW71609;
XX
DT 14-DEC-1998 (first entry)
XX
DE Agrobacterium sp. strain CP4 Class II EPSPS.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
KW N-phosphonomethylglycine.
XX
OS Agrobacterium sp.
XX
PN US5804425-A.
XX
PD 08-SEP-1998.
XX
PF 07-APR-1997; 97US-0833485.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
PR 07-APR-1997; 97US-0833485.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1998-505657/43.
DR N-PSDB; AAV58009.
XX
PT Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -
PT useful for characterisation of the enzyme to determine inhibition
PT data values
XX
PS Claim 3; Fig 3; 152pp; English.
XX

```
CC   An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
CC   sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
CC   -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
CC   S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
CC   Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
CC   produced by recombinant methods, can be used in kinetic studies to
CC   determine Ki and Km values of the enzyme for its characterisation. The
CC   enzyme is normally used for the production of 5-enolpyruvyl-3-
CC   phosphoshikimic acid in plants, and most forms of the enzyme are
CC   inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
CC   Inhibition data enables more accurate values of concentrations of
CC   herbicide to be used when growing the plant without being detrimental to
CC   it. This enables the plant to be grown in the presence of the herbicide,
CC   being used to inhibit the growth of undesired plants. The present
CC   sequence represents a Class II EPSPS from bacterial isolate
CC   Agrobacterium sp. strain CP4.
XX
SQ   Sequence   455 AA;

  Query Match             100.0%;  Score 2288;  DB 19;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-185;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   5
AAE05053
ID   AAE05053 standard; Protein; 455 AA.
XX
AC   AAE05053;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Agrobacterium sp. strain CP4 class II EPSPS protein.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   transformed bacteria; transgenic plant; herbicide.
XX
OS   Agrobacterium sp. CP4.
XX
PN   US6248876-B1.
XX
PD   19-JUN-2001.
XX
PF   20-AUG-1998;   98US-0137440.
XX
PR   13-SEP-1994;   94US-0306063.
PR   07-APR-1997;   97US-0833485.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 2001-407326/43.
DR   N-PSDB; AAD09754.
XX
PT   DNA probe capable of use in a polymerase chain reaction for identifying
PT   the presence of a target genomic DNA encoding a
PT   5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) enzyme -
XX
PS   Claim 3; Fig 3; 152pp; English.
XX
CC   The present invention relates to a DNA probe capable of use in a
CC   polymerase chain reaction for identifying the presence of a target
CC   genomic DNA encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS)
CC   enzyme referred to as class II EPSPS enzyme which is tolerant to
CC   glyphosate. EPSPS genes are useful in producing transformed bacteria
CC   and transgenic plants which are tolerant to glyphosate herbicide. The
CC   probe is useful for identifying the presence of a target genomic DNA
CC   encoding a EPSPS enzyme. The present sequence is Agrobacterium sp.
CC   strain CP4 class II EPSPS protein.
CC   Note: The present sequence, SED ID NO: 3 is stated as amino acid
CC   sequence throughout the specification. However, it is referred as probe
CC   in claim 3 of the specification.
XX
SQ   Sequence   455 AA;

  Query Match             100.0%;  Score 2288;  DB 22;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-185;
  Matches  455;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mshgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
```

```
Db    421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   6
AAR26449
ID   AAR26449 standard; Protein; 455 AA.
XX
AC   AAR26449;
XX
DT   28-JAN-1993  (first entry)
XX
DE   CP4-EPSPS.
XX
KW   pMON11030; CTP2; CP4; EPSPS; chloroplast transit peptide;
KW   5-enolpyruvyl-3-shikimate phosphate synthetase; ACC; ethylene;
KW   Pseudomonas chloroaphis 6G5.
XX
OS   Synthetic.
XX
WO9212249-A.
XX
PD   23-JUL-1992.
XX
PF   17-DEC-1991;   91WO-US09437.
XX
PR   26-DEC-1990;   90US-0632440.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Kishore GM,  Klee HJ;
XX
DR   WPI; 1992-284334/34.
DR   N-PSDB; AAQ27201.
XX
PT   Delaying fruit ripening and senescence in plants - by controlling
PT   ethylene prodn., pref. by expression of
PT   1-amino:cyclopropane-1-carboxylic acid deaminase
XX
PS   Disclosure; Page 64-66; 110pp; English.
XX
CC   The sequences given in AAR26448 and AAR26449 are encoded by genes which
CC   were used in the construction of an expression plasmid pMON11030 which
CC   was used to transform petunia plants.  This plasmid contained the genes
CC   for chloroplast transit peptide (CTP2) (AAR26448) and the CP4 synthetic
CC   5-enolpyruvyl-3-shikimate phosphate synthetase (EPSPS) gene (AAR26449)
CC   which is capable of conferring resistance to glyphosate.  The plasmid
CC   also contained the 1-aminocyclo-propane-1-carboxylic acid (ACC)
CC   deaminase gene from Pseudomonas chloroaphis (see also AAQ27199).  The
CC   transformed plants had ethylene levels reduced to about one half that
CC   of the control, untransformed plants.  It is expected that such plants
CC   will show reduced senesence of flowers and leaves when compared to
CC   untransformed plants.
XX
SQ   Sequence   455 AA;


  Query Match             99.7%;  Score 2282;  DB 13;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 2e-184;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455


RESULT   7
AAW39426
ID   AAW39426 standard; Protein; 455 AA.
XX
AC   AAW39426;
XX
DT   19-MAY-1998  (first entry)
XX
DE   CP4-EPSPS protein.
XX
KW   1-aminocyclopropane-1-carboxylic acid deaminase; ACC deaminase;
KW   fruit ripening; ethylene production; glyphosate resistance;
KW   5-enolpyruvyl-3-shikimate phosphate synthase; EPSPS; CP4.
XX
OS   Synthetic.
XX
PN   US5702933-A.
XX
PD   30-DEC-1997.
XX
PF   06-NOV-1995;   95US-0553943.
XX
PR   17-DEC-1991;   91US-0809457.
PR   26-DEC-1990;   90US-0632440.
PR   06-NOV-1995;   95US-0553943.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Kishore GM,  Klee HJ;
XX
DR   WPI; 1998-076419/07.
DR   N-PSDB; AAV09719.
XX
PT   Production of plants with delayed ripening - using DNA encoding
PT   1-amino:cyclo:propane-1-carboxylic acid deaminase
XX
PS   Example 9; Fig 21; 56pp; English.
XX
CC   This sequence represents the synthetic CP4 5-enolpyruvyl-3-shikimate
CC   phosphate synthase (EPSPS) gene which is capable of conferring
CC   resistance to glyphosate and is used in a novel method for producing
CC   fruit-bearing plants with delayed ripening. The method involves the
CC   expression of a 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase
CC   gene in a plant at a level sufficient to reduce ethylene production in
CC   the fruit.
XX
SQ   Sequence   455 AA;


  Query Match             99.7%;  Score 2282;  DB 19;  Length 455;
  Best Local Similarity   99.8%;  Pred. No. 2e-184;
  Matches  454;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 420

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     421 vtvddatmiatsfpefmdlmaglgakielsdtkaa 455
```

RESULT 8
AAM52214
ID AAM52214 standard; Protein; 527 AA.
XX
AC AAM52214;
XX
DT 12-FEB-2002 (first entry)
XX
DE EPSPS SEQ ID NO 3.
XX
KW Transgenic plant; herbicide resistance; EPSPS; GOX; Petunia hybrida;
KW 5-enolpyruvylshikimate-3-phosphate synthase; glyphosate oxidoreductase;
KW protoporphyrinogen IX oxidase; immunoglobulin; Agrobacterium; strain CP4;
KW chloroplast transit peptide.
XX
OS Petunia hybrida.
OS Agrobacterium sp.
XX
PN JP2001190168-A.
XX
PD 17-JUL-2001.
XX
PF 27-OCT-2000; 2000JP-0328811.
XX
PR 29-OCT-1999; 99JP-0310244.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
DR WPI; 2001-605307/69.
DR N-PSDB; ABA02854.
XX
PT New herbicide-resistant plant -
XX
PS Examples; Page 46-48; 66pp; Japanese.
XX
CC The invention relates to a transgenic plant which shows resistance to a
CC herbicide in an amount inhibiting natural
CC 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) activity of the plant
CC and having at least one enzymatic activity selected from:
CC (1) EPSPS activity different from natural EPSPS activity of the plant or
CC (2) glyphosate oxidoreductase (GOX) activity different from the natural
CC GOX activity of the plant; and in which a gene encoding a protein having
CC the following properties: (a) combines specifically with a substance
CC participating to the herbicidic activity a herbicide of
CC protoporphyrinogen IX oxidase inhibiting type; (b) has substantially no
CC denaturing activity on a substance to which said protein combines
CC specifically; and (c) contains substantially no framework region of the
CC variable region of immunoglobulin. The present sequence is that of the
CC Petunia hybrida EPSPS chloroplast transit peptide and the Agrobacterium
CC sp. strain CP4 EPSPS gene.
XX
SQ Sequence 527 AA;

Query Match 99.7%; Score 2282; DB 22; Length 527;
Best Local Similarity 99.8%; Pred. No. 2.5e-184;
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      73 mlhgassrpatarkssglsgtvripgdksishrsfmfgglasgetritgllegedvintg 132

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     133 kamqamgarirkegdtwiidgvgnggllapeapldfgnaatgcrltmglvgvydfdstfi 192

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     193 gdasltkrpmgrvlnplremgvqvksedgdrlpvtlrgpktptpityrvpmasaqvksav 252

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     253 llaglntpgittviepimtrdhtekmlqgfganltvetdadgvrtirlegrgkltgqvid 312

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     313 vpgdpsstafplvaallvpgsdvtilnvlmnptrtgliltlqemgadievinprlagged 372

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     373 vadlrvrsstlkgvtvpedrapsmideypilavaaafaegatvmngleelrvkesdrlsa 432

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     433 vanglklngvdcdegetslvvrgrpdgkglgnasgaavathldhriamsflvmglvsenp 492

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELSDTKAA 455
           |||||||||||||||||||||||||||||||||||
Db     493 vtvddatmiatsfpefmdlmaglgakielsdtkaa 527
```

RESULT 9
AAR22301
ID AAR22301 standard; Protein; 449 AA.
XX
AC AAR22301;
XX
DT 03-AUG-1992 (first entry)
XX
DE Class II EPSPS enzyme.
XX
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX
OS Achromobacter sp. strain LBAA.
XX
PN WO9204449-A.
XX
PD 19-MAR-1992.
XX
PF 28-AUG-1991; 91WO-US06148.
XX
PR 31-AUG-1990; 90US-0576537.

XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 5; 148pp; English.
XX
CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC phosphate synthase enzyme (EPSPS) of Achromobacter sp. strain LBAA
CC It is used to create glyphosate resistant plants or seeds which
CC can be planted in a field of crops to selectively control weeds.
CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC tobacco and alfalfa. This provides a cost effective, environmentally
CC compatible weed control device. See also AAR22300 and AAR22302.
XX
SQ Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 13; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.1e-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :|||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   | |||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 10
AAR22302
ID AAR22302 standard; Protein; 449 AA.
XX
AC AAR22302;
XX
DT 03-AUG-1992 (first entry)
XX
DE Class II EPSPS enzyme.
XX
KW Glyphosate tolerant bacteria; herbicides; glyphosate resistant plants;
KW 5'-enolpyruvylshikimate-3-phosphate synthase; weed control.
XX
OS Pseudomonas sp. strain PG2982.
XX
PN WO9204449-A.
XX
PD 19-MAR-1992.
XX
PF 28-AUG-1991; 91WO-US06148.
XX
PR 31-AUG-1990; 90US-0576537.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR;
XX
DR WPI; 1992-114356/14.
XX
PT DNA encoding class II 5'-enol pyruvyl shikimate-3-phosphate
PT synthase - for producing plants and bacteria tolerant to
PT glyphosate herbicides
XX
PS Disclosure; Fig 7; 148pp; English.
XX
CC The sequence is that of the Class II 5'-enolpyruvylshikimate-3
CC phosphate synthase enzyme (EPSPS) of Pseudomonas sp. strain PG2982
CC It is used to create glyphosate resistant plants or seeds which
CC can be planted in a field of crops to selectively control weeds.
CC The crops selected for are e.g. corn, wheat, rice, oilseed rape,
CC tobacco and alfalfa. This provides a cost effective, environmentally
CC compatible weed control device. See also AAR22300 and AAR22301.
XX
SQ Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 13; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.1e-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| ||||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :|||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   | |||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 11
AAW34684

```
ID   AAW34684 standard; Protein; 449 AA.
XX
AC   AAW34684;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Achromobacter sp. strain LBAA.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Achromobacter sp. strain LBAA.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          271..274
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93789.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Disclosure; Fig 4; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence was isolated from
CC   Achromobacter sp. strain LBAA. The EPSPS enzymes are part of the
CC   shikimic acid pathway, which leads to the biosynthesis of aromatic
CC   compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and
CC   3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC   inhibited by the herbicide glyphosate. It would be useful to produce
CC   transgenic crops containing glyphosate resistance genes so that
CC   glyphosate-containing herbicides can be applied to selectively kill
CC   weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for
CC   glyphosate, such that when introduced into a plant, the plant is made
CC   glyphosate-tolerant, and EPSPS enzyme activity is not affected. These
CC   class II EPSPS enzymes are fused to a chloroplast transit peptide to
CC   target the protein into the chloroplast, which is the site for the
CC   shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant
CC   under the control of a promoter such as figwort mosaic virus promoter or
CC   the cauliflower mosaic virus promoter, so that expression is enhanced.
XX
SQ   Sequence  449 AA;
```

```
  Query Match             83.3%;  Score 1906.5;  DB 18;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 1.1e-152;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :|||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           .:|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:|||||:|||||||||||||:||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||||| |||||| |||||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

```
RESULT  12
AAW34685
ID   AAW34685 standard; Protein; 449 AA.
XX
AC   AAW34685;
XX
DT   17-FEB-1998  (first entry)
XX
DE   Class II EPSP synthase (EPSPS) from Pseudomonas sp. strain PG2982.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; glyphosate;
KW   shikimic acid pathway; phosphoenolpyruvic acid; 3-phosphoshikimic acid;
KW   5-enolpyruvyl-3-phosphoshikimic acid; herbicide; transgenic plant;
KW   glyphosate resistance gene; glyphosate-tolerance; promoter.
XX
OS   Pseudomonas sp. strain PG2982.
XX
FH   Key             Location/Qualifiers
FT   Region          200..204
FT                   /label= characteristic_region
FT                   /note= "see AAW34690"
FT   Region          26..29
FT                   /label= characteristic_region
FT                   /note= "see AAW34691"
FT   Region          173..177
FT                   /label= characteristic_region
FT                   /note= "see AAW34692"
FT   Region          273..276
FT                   /label= characteristic_region
FT                   /note= "see AAW34693"
XX
PN   US5627061-A.
XX
PD   06-MAY-1997.
```

```
XX
PF   07-JUN-1995;   95US-0476008.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-271315/24.
DR   N-PSDB; AAT93790.
XX
PT   Production of glyphosate-herbicide tolerant plants - using DNA
PT   encoding class II 5-enol:pyruvyl:shikimate-3-phosphate synthase
PT   enzyme(s)
XX
PS   Disclosure; Fig 5; 151pp; English.
XX
CC   AAW34683-89 represent a new class of glyphosate-tolerant
CC   5-enolpyruvylshikimate-3-phosphate synthases (EPSPS). These novel
CC   EPSPS enzymes have little homology with known Class I EPSPS enzymes, and
CC   belong to a new class, Class II. The present sequence was isolated from
CC   Pseudomonas sp. strain PG2982. The EPSPS enzymes are part of the
CC   shikimic acid pathway, which leads to the biosynthesis of aromatic
CC   compounds. EPSPS converts phosphoenolpyruvic acid (PEP) and
CC   3-phosphoshikimic acid to 5-enolpyruvyl-3-phosphoshikimic acid, and is
CC   inhibited by the herbicide glyphosate. It would be useful to produce
CC   transgenic crops containing glyphosate resistance genes so that
CC   glyphosate-containing herbicides can be applied to selectively kill
CC   weeds. The novel EPSPS enzymes exhibit a low Km for PEP and a high Ki for
CC   glyphosate, such that when introduced into a plant, the plant is made
CC   glyphosate-tolerant, and EPSPS enzyme activity is not affected. These
CC   class II EPSPS enzymes are fused to a chloroplast transit peptide to
CC   target the protein into the chloroplast, which is the site for the
CC   shikimic acid pathway. In addition, the EPSPS gene is cloned into a plant
CC   under the control of a promoter such as figwort mosaic virus promoter or
CC   the cauliflower mosaic virus promoter, so that expression is enhanced.
XX
SQ   Sequence   449 AA;

  Query Match             83.3%;  Score 1906.5;  DB 18;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 1.1e-152;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||||  || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||:: |||||:|:|| ||||  ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || ||||||||||   |  |||||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447


RESULT  13
AAW24479
ID   AAW24479 standard; Protein; 449 AA.
XX
AC   AAW24479;
XX
DT   02-OCT-1997  (first entry)
XX
DE   Class II EPSPS used for glyphosate resistant plant production.
XX
KW   5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW   glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW   fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW   tobacco.
XX
OS   Achromobacter sp. strain LBAA.
XX
PN   US5633435-A.
XX
PD   27-MAY-1997.
XX
PF   31-AUG-1990;   90US-0576537.
XX
PR   13-SEP-1994;   94US-0306063.
PR   31-AUG-1990;   90US-0576537.
PR   28-AUG-1991;   91US-0749611.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Barry GF,  Kishore GM,  Padgette SR,  Stallings WC;
XX
DR   WPI; 1997-297418/27.
DR   N-PSDB; AAT77315.
XX
PT   New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT   used for transforming plants to produce plants which are tolerant to
PT   glyphosate herbicide
XX
PS   Claim 7; Column 63-66; 154pp; English.
XX
CC   AAW24479 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC   phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC   to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC   the production of herbicide resistant (glyphosate-tolerant) plants
CC   such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC   rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC   pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ   Sequence   449 AA;

  Query Match             83.3%;  Score 1906.5;  DB 18;  Length 449;
  Best Local Similarity   83.1%;  Pred. No. 1.1e-152;
  Matches  374;  Conservative   32;  Mismatches   41;  Indels    3;  Gaps    1;

Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:|||||  || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:|||||||||||||:: |||||:|:|| ||||  ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180
```

```
Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:||||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||||| |||||| |||||||||||||||||||||||||:| |||||||
Db    241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||||:|
Db    301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   |  ||||||||||||||||||| :| |
Db    361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: |||||||||||:| |||||||||
Db    418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 14
AAW24480
ID AAW24480 standard; Protein; 449 AA.
XX
AC AAW24480;
XX
DT 02-OCT-1997 (first entry)
XX
DE Class II EPSPS used for glyphosate resistant plant production.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; EPSPS; Class II;
KW glyphosate resistant; transgenic plant; herbicide; shikimic acid;
KW fusion protein; soybean; corn; wheat; barley; cotton; sugarbeet;
KW tobacco.
XX
OS Pseudomonas sp. strain PG2982.
XX
PN US5633435-A.
XX
PD 27-MAY-1997.
XX
PF 31-AUG-1990; 90US-0576537.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1997-297418/27.
DR N-PSDB; AAT77316.
XX
PT New isolated 5-enol:pyruvyl:shikimate-3-phosphate synthase gene -
PT used for transforming plants to produce plants which are tolerant to
PT glyphosate herbicide
XX
PS Claim 7; Column 69-72; 154pp; English.
XX
CC AAW24480 shows the sequence of a class II 5-enolpyruvylshikimate-3-
CC phosphate synthase (EPSPS) enzyme. Class II EPSPS enzymes are tolerant
CC to glyphosate herbicides. EPSPS and sequences encoding it are used for
CC the production of herbicide resistant (glyphosate-tolerant) plants
CC such as corn, wheat, rice, barley, soybean, cotton, sugarbeet, oilseed
CC rape, canola, flax, sunflower, potato, tobacco, tomato, alfalfa, poplar,
CC pine, eucalyptus, apple, lettuce, peas, lentils, grape and turf grasses.
XX
SQ Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 18; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.1e-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```
Qy      1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
          ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db      1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy     61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
          :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db     61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy    121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
          |||||:|||||||||||||||:: ||||:|:|| |||| ||||||||||||||||
Db    121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy    181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
          |||||||||:||||||:||||||||||||||||:|||||| |||| ||: |:||| || ||
Db    181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy    241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
          |||||||||||||||||| |||||| |||||||||||||||||||||||||:| |||||||
Db    241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy    301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
          |||||||:| |||| || :||||||||||:||:||:|||| |||:||:|||||||||||:|
Db    301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy    361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
          || ||: ||||| ||| || |||||||||||   |  ||||||||||||||||||| :| |
Db    361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy    421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
          |||||: |||||||||||:| |||||||||
Db    418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

RESULT 15
AAW71611
ID AAW71611 standard; Protein; 449 AA.
XX
AC AAW71611;
XX
DT 14-DEC-1998 (first entry)
XX
DE Pseudomonas sp. strain PG2982 Class II EPSPS.
XX
KW 5-enolpyruvylshikimate-3-phosphate synthase; class II EPSPS enzyme;
KW glyphosate herbicide; transformed bacteria; class I EPSPS enzyme;
KW resistance; inhibition; 5-enolpyruvyl-3-phosphoshikimic acid;
KW N-phosphonomethylglycine.
XX
OS Pseudomonas sp.
XX
PN US5804425-A.
XX
PD 08-SEP-1998.
XX
PF 07-APR-1997; 97US-0833485.
XX
PR 13-SEP-1994; 94US-0306063.
PR 31-AUG-1990; 90US-0576537.
PR 28-AUG-1991; 91US-0749611.
PR 07-APR-1997; 97US-0833485.
XX
PA (MONS) MONSANTO CO.
XX
PI Barry GF, Kishore GM, Padgette SR, Stallings WC;
XX
DR WPI; 1998-505657/43.
DR N-PSDB; AAV58012.
XX
PT Glyphosate resistant 5-enol:pyruvyl:shikimate-3-phosphate synthase -

PT useful for characterisation of the enzyme to determine inhibition
PT data values
XX
PS Disclosure; Fig 5; 152pp; English.
XX
CC An isolated 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) has the
CC sequence domains (I), (II), (III) and (IV): -R-X1-H-X2-E- (I),
CC -G-D-K-X3- (II); -S-A-Q-X4-K- (III); and -N-X5-T-R- (IV). Where X1 = G,
CC S, T, C, Y, N, Q, D or E; X2 and X3 = S or T; X4 and X5 = A, R, N, D, C,
CC Q, E, G, H, I, L, K, M, F, P, S, T, W, Y or V. The EPSPS enzyme,
CC produced by recombinant methods, can be used in kinetic studies to
CC determine Ki and Km values of the enzyme for its characterisation. The
CC enzyme is normally used for the production of 5-enolpyruvyl-3-
CC phosphoshikimic acid in plants, and most forms of the enzyme are
CC inhibited by N-phosphonomethylglycine ('glyphosate') herbicides.
CC Inhibition data enables more accurate values of concentrations of
CC herbicide to be used when growing the plant without being detrimental to
CC it. This enables the plant to be grown in the presence of the herbicide,
CC being used to inhibit the growth of undesired plants. The present
CC sequence represents a Class II EPSPS from bacterial isolate
CC Pseudomonas sp. strain PG2982.
XX
SQ Sequence 449 AA;

Query Match 83.3%; Score 1906.5; DB 19; Length 449;
Best Local Similarity 83.1%; Pred. No. 1.1e-152;
Matches 374; Conservative 32; Mismatches 41; Indels 3; Gaps 1;

```
Qy       1 MSHGASSRPATARKSSGLSGTVRIPGDKSISHRSFMFGGLASGETRITGLLEGEDVINTG 60
           ||| || :|||||:|  |:| :||||||||||||||||||||||||||||||||||||||
Db       1 mshsaspkpatarrsealtgeiripgdksishrsfmfgglasgetritgllegedvintg 60

Qy      61 KAMQAMGARIRKEGDTWIIDGVGNGGLLAPEAPLDFGNAATGCRLTMGLVGVYDFDSTFI 120
           :|||||||:|||||| |||:||||| || ||| |||||| || |||||||| ||  ::||
Db      61 ramqamgakirkegdvwiingvgngcllqpeaaldfgnagtgarltmglvgtydmktsfi 120

Qy     121 GDASLTKRPMGRVLNPLREMGVQVKSEDGDRLPVTLRGPKTPTPITYRVPMASAQVKSAV 180
           |||||:||||||||||||||:: |||||:|:|| |||| ||||||||||||||||||
Db     121 gdaslskrpmgrvlnplremgvqveaadgdrmpltligpktanpityrvpmasaqvksav 180

Qy     181 LLAGLNTPGITTVIEPIMTRDHTEKMLQGFGANLTVETDADGVRTIRLEGRGKLTGQVID 240
           |||||||||:||||||:|||||||||||||||:|||||| |||| ||: |:||| || ||
Db     181 llaglntpgvttviepvmtrdhtekmlqgfgadltvetdkdgvrhiritgqgklvgqtid 240

Qy     241 VPGDPSSTAFPLVAALLVPGSDVTILNVLMNPTRTGLILTLQEMGADIEVINPRLAGGED 300
           |||||||||||||||||| |||||| |||||||||||||||||||||:| |||||||
Db     241 vpgdpsstafplvaallvegsdvtirnvlmnptrtgliltlqemgadievlnarlagged 300

Qy     301 VADLRVRSSTLKGVTVPEDRAPSMIDEYPILAVAAAFAEGATVMNGLEELRVKESDRLSA 360
           |||||||:| |||| || :||||||||||:||:||:|||| |||:||:||||||||||:|
Db     301 vadlrvrasklkgvvvpperapsmideypvlaiaasfaegetvmdgldelrvkesdrlaa 360

Qy     361 VANGLKLNGVDCDEGETSLVVRGRPDGKGLGNASGAAVATHLDHRIAMSFLVMGLVSENP 420
           || ||: ||||| ||| || |||||||||||   | |||||||||||||||||| :| |
Db     361 vargleangvdctegemsltvrgrpdgkglg---ggtvathldhriamsflvmglaaekp 417

Qy     421 VTVDDATMIATSFPEFMDLMAGLGAKIELS 450
           |||||: ||||||||||:| |||||||||
Db     418 vtvddsnmiatsfpefmdmmpglgakiels 447
```

Search completed: August 15, 2002, 13:58:47
Job time: 66 sec